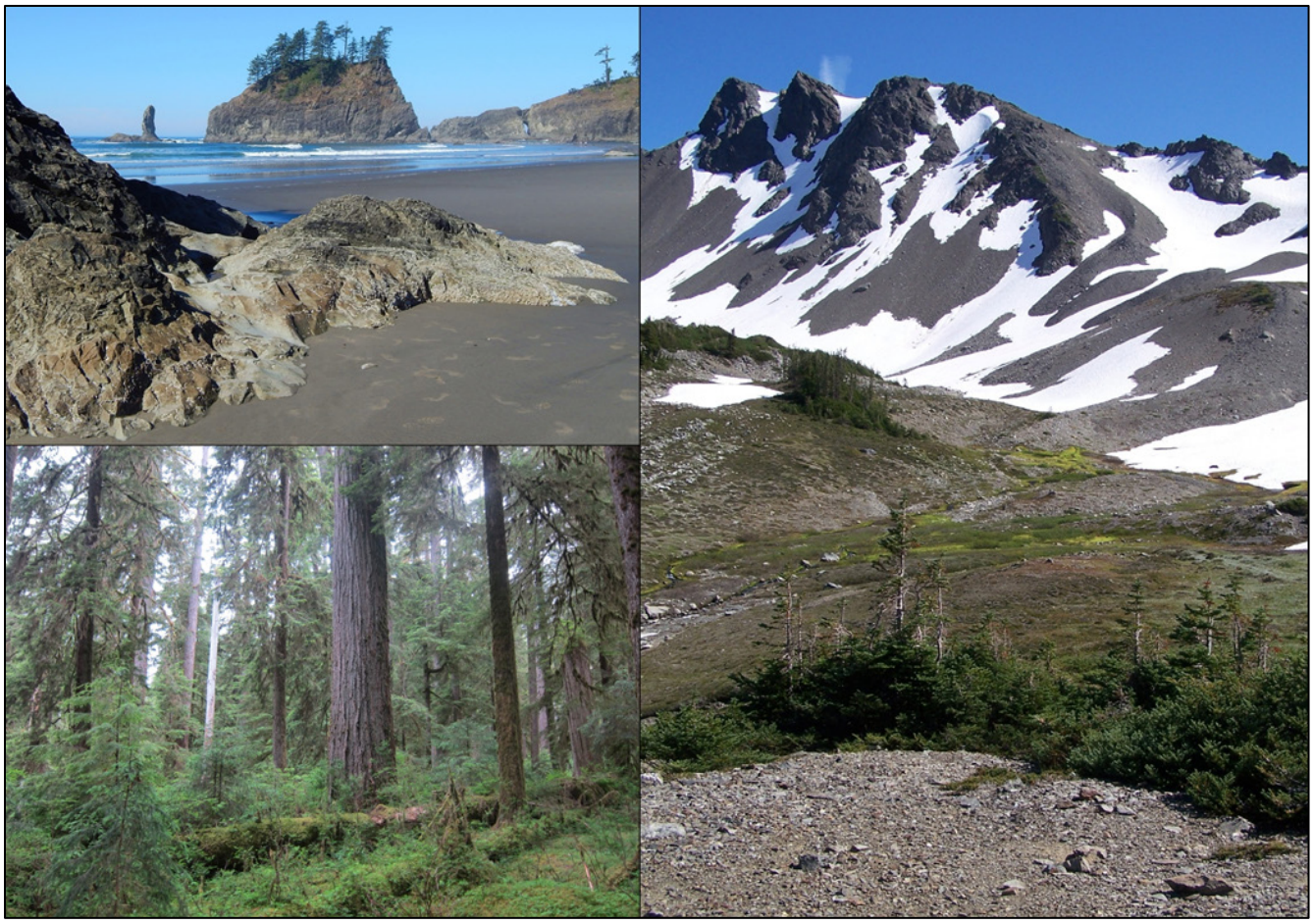
National Park Service
U.S. Department of the Interior

Natural Resource Stewardship and Science

Olympic National Park Vegetation Classification and Mapping Project Report

Natural Resource Report NPS/NCCN/NRR—2021/2255



ON THIS PAGE
An NPS field crew takes advantage of a sun break on the Olympic coast.
Image credit NPS

ON THE COVER
Three worlds of Olympic National Park: coastal headlands, temperate forests, and alpine environments.
Image credit NPS

Olympic National Park Vegetation Classification and Mapping Project Report

Natural Resource Report NPS/NCCN/NRR—2021/2255

Eric M. Nielsen,[1] Catharine Copass,[2] Rachel L. Brunner,[1] Lindsey K. Wise[1]

[1]Institute for Natural Resources
Portland State University
Portland, Oregon

[2]National Park Service
Olympic National Park
Port Angeles, Washington

May 2021

U.S. Department of the Interior
National Park Service
Natural Resource Stewardship and Science
Fort Collins, Colorado

The National Park Service, Natural Resource Stewardship and Science office in Fort Collins, Colorado, publishes a range of reports that address natural resource topics. These reports are of interest and applicability to a broad audience in the National Park Service and others in natural resource management, including scientists, conservation and environmental constituencies, and the public.

The Natural Resource Report Series is used to disseminate comprehensive information and analysis about natural resources and related topics concerning lands managed by the National Park Service. The series supports the advancement of science, informed decision-making, and the achievement of the National Park Service mission. The series also provides a forum for presenting more lengthy results that may not be accepted by publications with page limitations.

All manuscripts in the series receive the appropriate level of peer review to ensure that the information is scientifically credible, technically accurate, appropriately written for the intended audience, and designed and published in a professional manner.

This report received informal peer review by subject-matter experts who were not directly involved in the collection, analysis, or reporting of the data.

Views, statements, findings, conclusions, recommendations, and data in this report do not necessarily reflect views and policies of the National Park Service, U.S. Department of the Interior. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. Government.

This report is available in digital format from the NPS North Coast and Cascades Network website and the Natural Resource Publications Management website. If you have difficulty accessing information in this publication, particularly if using assistive technology, please email irma@nps.gov.

Please cite this publication as:

Nielsen, E. M., C. Copass, R. L. Brunner, and L. K. Wise. 2021. Olympic National Park vegetation classification and mapping project report. Natural Resource Report NPS/NCCN/NRR—2021/2255. National Park Service, Fort Collins, Colorado. https://doi.org/10.36967/nrr-2286420.

NPS 149/175302, May 2021

Contents

Contents (continued)

Contents (continued)

Contents (continued)

Figures

Page

Tables

Page

Supplements

Available as separate PDFs through the NPS Data Store at https://irma.nps.gov, or via other media upon request to the NPS North Coast and Cascades Network Data Manager.

Supplement 1: Vegetation map

Nielsen, E. M., C. Copass, R. L. Brunner, and K. Braun. 2021. Olympic National Park vegetation map. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Supplement 2: Map class descriptions

Nielsen, E. M., R. L. Brunner, C. Copass, and L. K. Wise. 2021. Olympic National Park map class descriptions. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Supplement 3: Map class key

Brunner, R. L., C. Copass, L. K. Wise, and E. M. Nielsen. 2021. Olympic National Park map class key. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Supplement 4: Accuracy assessment contingency tables

Institute for Natural Resources. 2021. Olympic National Park accuracy assessment contingency tables. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Supplement 5: Taxonomic tables

Institute for Natural Resources. 2021. Taxonomic tables for Mount Rainier, Olympic, and North Cascades National Parks. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2283943.

Supplement 6: Mapping association descriptions

Nielsen, E. M., and R. L. Brunner. 2021. Vegetation associations for mapping Pacific Northwest national parks. Institute for Natural Resources, Portland State University, Portland, Oregon. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2283945 (available June 2021).

Executive Summary

The Vegetation Mapping Inventory (VMI) is an effort by the National Park Service (NPS) to classify, describe, and map vegetation communities present on NPS units across the United States. The Institute for Natural Resources, working in cooperation with the NPS North Coast and Cascades Network (NCCN), has completed a VMI project for the vegetation communities of Olympic National Park (OLYM).

The map is based on a vegetation classification developed during the project and was created using an inductive modeling approach. Data used to construct the classification were collected between 2005 and 2015 and included plots from Mount Rainier National Park and North Cascades National Park Complex. These plots were used to develop and refine the association-level National Vegetation Classification (NVC). The associations were combined into map classes based roughly on the NVC alliance-level classification, but updated to allow improved map detail and accuracy. Model training data relied only on plots from OLYM, collected between 2005 and 2019. Independent field accuracy assessment data were collected in 2013 and applied to the final map generated later.

The map development process was organized around the random forests machine learning algorithm. The modeling used 2,519 plots representing 150 vegetation associations and 50 map classes. Imagery from the National Agriculture Imagery Program and the Sentinel-2 and Landsat 8 satellites, airborne lidar bare earth and canopy height data, elevation data from the U.S. Geological Survey 3D Elevation Program, and climate normals from the PRISM Climate Group were used to develop a variety of predictor metrics. The predictors and the map class calls at each plot were input to a process in which each map class was modeled against every other map class in a factorial random forests scheme. We used the plot-level modeling outcomes and species composition data to adjust the crosswalk between association and map class so that floristic consistency and model accuracy were jointly optimized across all classes. The map was produced by predicting the factorial models and selecting the overall best-performing class at each 3-meter pixel.

The final vegetation map, including a buffer surrounding the park, contains 43 natural vegetated classes, seven mostly unvegetated natural classes, and four classes representing burned areas or anthropogenic disturbance. Coniferous forests and woodlands cover four-fifths of the park, with nearly half that consisting of forests dominated by silver fir (*Abies amabilis*). Upper montane and subalpine forests and woodlands cover one-fifth, mostly surrounding Mount Olympus and on the east side of the park. Fourteen percent of the park is occupied by wet lowland forests in west-side valleys and on the coastal strip. Drier lowland forests are found mainly on the north and east sides and account for another thirteen percent. Broadleaf and mixed forests occupy less than three percent of the park, mostly in west-side river floodplains. Shrub-dominated and herbaceous vegetation each cover about four percent, mostly in avalanche tracks and as high-elevation dwarf shrublands and dry meadows. Sparsely vegetated and entirely bare rock cover six percent of the park; lake surfaces and exposed snow and ice round out most of the remaining three percent.

The accuracy assessment was based on 859 independent field-collected plots representing all the vegetated classes, as well as alluvial and colluvial barrens, which also often host vegetation

communities. They were gathered from an inference area covering 6.5% of the park. The overall map accuracy based on this sample was 86.0%. After correcting for map class prevalence in the inference area, the overall accuracy was 84.1%. Eleven of the 45 classes evaluated in the AA failed to meet the 80% NPS standard for user's accuracy; ten fell short of the standard for producer's accuracy. The AA discussion in the report contains a review of all classes failing to meet either standard, considers possible remedies for each, and provides recommendations to NPS for possible modifications of the map in response to the issues identified.

Many new methodologies for mapping and floristic analysis were developed during this project. These innovations were also applied in mapping the other large NCCN national parks. In addition to allowing the development of this series of maps, these methods should be useful to the NCCN and VMI for other mapping projects and purposes. Products resulting from this project include (a) this report, (b) the report supplements listed above, (c) a geodatabase with map polygon attributes, plot locations, and project boundaries, (d) training and accuracy assessment plot field forms and data, including ground photography, (e) hard copy vegetation maps and (f) metadata for digital products. Geospatial products are provided in the Universal Transverse Mercator (UTM) Zone 10 projection using the North American Datum of 1983.

Acknowledgments

Many skilled individuals came together to complete this project over its fifteen-year duration. We are especially grateful to John Boetsch, Katherine Braun, Mark Huff, Matt Lee, Stacy McDonough and Tynan Ramm-Granberg, whose efforts over many years were truly instrumental to the project's success. The names and roles of all contributors are listed below; we sincerely apologize to anyone we have overlooked.

National Park Service

Steve Acker, OLYM supervisory botanist — project support, report peer review
Katherine Braun, OLYM GIS specialist — production of map sheets for mapping and accuracy assessment (AA) fieldwork, production of final maps, GPS and metadata assistance
Bill Bennett — mapping crew member
John Boetsch, network data manager/ecologist — creation and maintenance of field plot database
Karl Brown, national VIP manager — programmatic guidance
Erin Burke — mapping crew member
Joshua Chenoweth, OLYM restoration biologist — classification crew leader
Janet Coles, OLYM supervisory botanist — project support, draft and final report review
Tammy Cook, biologist — programmatic guidance
Catharine Copass, network VIP projects manager — project oversight and coordination
Ida Cunningham — classification crew member, mapping crew member
Holly Faulstich — mapping crew member
Lise Grace, NCCN publication manager — report editing and publication assistance
Dawn Graham — classification crew member
Liz Graham — AA crew member
Scott Gremel, OLYM wildlife biologist — inputs to draft and final map projects, report peer review
Evan Hegarty — mapping crew member
Mark Huff, *NCCN Program Coordinator* — programmatic guidance, project support
Emilie Kohler — AA crew member
Allen McCoy, regional GIS specialist — data procurement
Stacy McDonough — mapping crew leader
Colin Meston — classification crew member
Tyler Morrison — classification crew member
Laurel Moulton — classification crew member
Evan Raimundo — mapping crew member
Tynan Ramm-Granberg — mapping crew member, mapping data QC assistance, vegetation classification assistance
Anya Tyson — AA crew member
Ben Wallace — mapping crew member
Marisa Whisman — classification crew leader

Institute for Natural Resources, Portland State University

Rachel Brunner, vegetation ecologist — field data QC, vegetation classification and image processing assistance, mapping crew member

John Christy, wetlands ecologist — vegetation classification assistance

Eleanor Gaines, INR-Portland director — project management assistance

Jimmy Kagan, ecologist — INR principal investigator (2008–11), vegetation classification assistance

Seth Keena-Levin — AA crew member

Matt Lee, field data manager — AA crew leader, field data QC lead, vegetation classification assistance, mapping crew member

Eric Nielsen, geospatial scientist — INR PI (2012–20), vegetation classification, image processing, modeling, statistical analysis, sample and field protocol designs, AA crew member

Lindsey Koepke Wise, biodiversity data manager — report editing, formatting and graphics production, vegetation taxonomy assistance

NatureServe

Gwen Kittel, vegetation ecologist — NVC alliance development for NCCN

Marion Reid, vegetation ecologist — NVC alliance development for NCCN

Funding for this work was provided by the I&M VIP program, through the Pacific Northwest Cooperative Ecosystem Study Unit, Task Agreements P10AC00407 (originally J8W07100012) and P15AC00990.

Acronyms and Abbreviations

3DEP	3D Elevation Program (USGS), provider of DEMs
AA	accuracy assessment (see following glossary)
AIA	attempted inference area (see following glossary)
CE	commission error (see following glossary)
DBH	diameter at breast height, the diameter of a tree measured at 4.5 feet
DEM	digital elevation model; digitally represented elevation data
DIT	differential indicators tool, developed at INR
EBLA	Ebey's Landing National Historical Reserve
FGDC	Federal Geographic Data Committee
GPS	Global Positioning System
HALT	hybrid assemblage labeling tool, developed at INR
I&M	Inventory and Monitoring Program (NPS)
INR	Institute for Natural Resources (Portland State University)
LEWI	Lewis and Clark National and State Historical Parks
MORA	Mount Rainier National Park
MMU	minimum mapping unit (see following glossary)
NAD83	North American Datum of 1983; the geographic datum for this project
NAIP	National Agriculture Imagery Program (USDA)
NCCN	NPS North Coast and Cascades Network
NOCA	North Cascades National Park
NPS	National Park Service
NTM	nested texture metrics (see following glossary)
NVC	National Vegetation Classification
NVCS	National Vegetation Classification Standard
OE	omission error (see following glossary)
OLI	Operational Land Imager (Landsat 8, 2013–present)
OLYM	Olympic National Park
PA	producer's accuracy (see following glossary)
PCT	population contingency table (see following glossary)
PI	photo-interpretation, use of imagery to ascertain land cover characteristics
PNV	potential natural vegetation (see following glossary)
PNW	Pacific Northwest region
PRISM	Parameter-elevation Regressions on Independent Slopes Model (PRISM Climate Group 2019)
PSU	Portland State University
QC	quality control (see following glossary)
RF	random forests (see following glossary)
RIA	reached inference area (see following glossary)
RMSE	root-mean-square error

RRRF	round robin random forests (see following glossary)
SAGA	System for Automated Geoscientific Analyses (Conrad et al. 2015)
SCM	species cover match tool (see following glossary)
SCT	sample contingency table (see following glossary)
TM	Thematic Mapper (Landsat 4 and 5, 1984–2011)
TNC	The Nature Conservancy
UA	user's accuracy (see following glossary)
USDA	United States Department of Agriculture
USFS	United States Forest Service
USGS	United States Geological Survey
UTM10	Universal Transverse Mercator zone 10; standard geographic projection for the project
UW	University of Washington
VIP	Vegetation Inventory Program (NPS)
VMI	Vegetation Mapping Inventory (NPS)
WNHP	Washington Natural Heritage Program

Glossary

accuracy assessment	Statistical analysis to determine the degree to which a map correctly represents on-the-ground conditions.
accuracy assessment plots	Field plots collected during the third major phase of field sampling, used as ground truth in the AA.
attempted inference area	The spatial region within which plots were targeted in the AA sample design.
classification plots	Field plots collected during the first major phase of field sampling, used to define the initial vegetation associations (Crawford et al. 2009). Supplemented by the mapping plots, they were also used to define the mapping associations and map classes, and to create model training data.
commission error	The frequency with which a map specifies the presence of a class where it is not actually present.
contingency table	An AA error matrix documenting the extent of class-specific confusion between mapped and ground-truth data (often called a "confusion matrix").
floristic similarity	The degree of species composition resemblance between two plots, associations or map classes.
full-ocular plot	A field sample including reasonably complete species cover data.
inductive model	A predictive representation of reality built from provided examples.
Landsat	Mid-resolution U.S. remote sensing satellites, active from 1972–present. The Landsat data used were at 30-meter resolution.
lidar	Light detection and ranging; a laser-based technology for measuring elevation.
map classes	The thematic units to which map polygons are labeled; formed by merging similar mapping associations.
mapping associations	The fundamental classification units on which the map classes and therefore the NCCN maps are based; formed by revising the NVC associations (Crawford et al. 2009, Ramm-Granberg et al. 2021) for increased floristic and modeling consistency.
mapping plots	Field plots collected during the second major phase of field sampling. Supplemented by the classification plots, they were used to define the mapping associations and map classes, and to create model training data.
minimum mapping unit	The smallest homogeneous area intended for representation in the map; for this project, nominally 500 square meters (0.05 hectares).
modeling similarity	The susceptibility of two plots, associations or map classes to incorrect labeling in inductive modeling; in other words, their degree of similarity in predictor data.

nested texture metrics	A method for extracting multi-resolution spatial patterning information from imagery, developed at INR.
omission error	The frequency with which a map neglects to show a class where it is actually present.
partial-ocular plot	A field sample with incomplete species cover data.
patch	A fairly homogeneous and contiguous area of land cover discernible on the ground, typically composed of a single vegetation or abiotic land cover type.
photointerpreted plots	Plots assigned to map class based on an assessment of imagery and other data sources available in the office, mostly used for training and AA of abiotic classes.
producer's accuracy	The estimated probability that a map is correct where a particular map class is found on the ground.
population contingency table	An AA error matrix scaled to the mapped extent of each class in the inference area.
potential natural vegetation	The vegetation type that would hypothetically exist at a location under a natural disturbance regime.
predictor data	Independent data (e.g., variables derived from imagery, topography, climate, etc.) provided to an inductive model for prediction of a dependent variable (e.g., a map class).
Python	The programming language used for most project geoprocessing.
quality control	The process of improving the quality of data collected and/or entered.
R	The programming language used for most project statistical analyses.
reached inference area	The portion of the AIA reached by AA field crews and from which accuracy conclusions were drawn.
random forests	An outlier-resistant inductive modeling algorithm (Breiman 2001).
round robin random forests	An extension to random forests developed at INR for modeling a large number of classes with reduced sample size-related bias.
sample contingency table	An AA error matrix based on raw numbers of samples.
SCM taxa	The botanical taxa on which floristic similarity between plots, associations and map classes was determined; mostly species, but also including some genera and some sub-genus groupings of species.
Sentinel-2	Mid-resolution European remote sensing satellites, active from 2015–present. The Sentinel-2 data used were at 10- to 20-meter resolution.
species cover match	A set of tools for evaluating the degree of fit between a plot and an association or map class. Variants were created for use with full-ocular and partial-ocular plots.
training data	Locations confidently assigned to a particular map class and used to build inductive models connecting patterns in predictor data to that map class.
user's accuracy	The estimated probability that a map is correct where a particular map class is mapped.

1. Introduction

1.1. Background

1.1.1. NPS Vegetation Mapping Inventory and National Vegetation Classification

The Vegetation Mapping Inventory (VMI) was created to classify, map and describe vegetation communities on National Park Service units across the United States (NPS 2018). The resulting classifications, maps and reports contribute to the inventory of NPS resources and inform management and planning decisions. NPS has provided guidelines for vegetation classification (Lea 2011) and map accuracy assessment (Lea and Curtis 2010).

VMI maps are based on the National Vegetation Classification (NVC), a collaborative effort to classify the vegetation communities of the U.S. in a consistent manner. The NVC grew out of work by The Nature Conservancy (TNC), NatureServe, and the Natural Heritage Program network (Grossman et al. 1998). It is an evolving classification to which several federal agencies and non-profit organizations—including NPS, the U.S. Fish and Wildlife Service, the U.S. Geological Survey, TNC, and the Ecological Society of America—have contributed.

The National Vegetation Classification Standard (NVCS) provides the hierarchical structure for the NVC. Based in part on an earlier international classification (UNESCO 1973), it was originally adopted by the Federal Geographic Data Committee (FGDC) in 1997 and updated substantially by FGDC (2008). The upper levels of the hierarchy define classes based on broad-scale physiognomic and ecological factors (e.g., climate regimes, continentality), the middle levels incorporate floristic and additional physiognomic factors based on finer scale variation, and the lower levels are based entirely on floristics, including dominant and diagnostic overstory and understory species. The hierarchy for natural vegetation and the classification for an association found in Pacific Northwest (PNW) montane forests is shown in **Table 1**. The most recent revision of the NVC was published as USNVC (2019).

Table 1. National Vegetation Classification System hierarchy (version 2, FGDC 2008), and names of all levels for an example association.

Hierarchy level	Name	Code
Level 1–Class	Forest & Woodland	C01
Level 2–Subclass	Temperate & Boreal Forest & Woodland	S15
Level 3–Formation	Cool Temperate Forest & Woodland	F008
Level 4–Division	Vancouverian Forest & Woodland	D192
Level 5–Macrogroup	Vancouverian Subalpine-High Montane Forest	M025
Level 6–Group	North-Central Pacific Mountain Hemlock-Silver Fir Woodland	G849
Level 7–Alliance	Tsuga mertensiana-Abies amabilis Forest & Woodland	A3723
Level 8–Association	Abies amabilis/Rhododendron albiflorum Forest	CEGL000225

1.1.2. NCCN vegetation inventory project

The North Coast and Cascades Network (NCCN) vegetation inventory project (VIP) began in 2005. The first several years were primarily devoted to developing the regional association-level NVC (Crawford et al. 2009). The Institute for Natural Resources (INR) joined the project in 2008 to assist with the vegetation mapping portion of the project. In addition to the large parks—Mount Rainier National Park (MORA, 956 km^2), Olympic National Park (OLYM, 3734 km^2) and the North Cascades National Park Complex (NOCA, 2769 km^2)—INR and NPS also worked cooperatively to complete two other mapping projects, the Lewis and Clark National and State Historical Parks (LEWI, 38 km^2; Kagan et al. 2012) and Ebey's Landing National Historical Reserve (EBLA, 78 km^2; Copass and Ramm-Granberg 2016a).

1.1.3. OLYM vegetation classification and mapping project

The three large NCCN parks were treated as a single mapping endeavor, but delivered as three distinct projects and reports (see also Nielsen et al. 2021a, Nielsen et al. 2021b). Although much of the classification and mapping work proceeded concurrently, the fieldwork focus moved from one park to another during the map training and accuracy assessment phases. The OLYM project was the second of the parks to be sampled in each phase, with training data collection from 2009 to 2012 and accuracy assessment fieldwork in 2013. The work at OLYM benefitted substantially from the development of a consistent field sampling strategy and the cultivation of an experienced field crew. Given the magnitude and complexity of the challenges posed by sampling this very large and diverse park, those advantages were key to the project's success.

1.2. Approach

1.2.1. Classification

Mountainous environments in the Pacific Northwest present interlocking challenges for vegetation classification. First, the environmental envelopes of most species are largely determined by local climate, which responds in spatially continuous[1] and often complex patterns with elevation, aspect, and characteristics of the surrounding terrain. Along these gradients, competition and historic factors may result in gradual changes in species prominence that make field assessment of breaks based on thresholds of species cover difficult; chance variation adds to that unreliability. Second, species succession is often drawn out over centuries, and its rate varies over both coarse and fine spatial scales. For example, in the montane zone, the most characteristic successional process is the gradual establishment and increase in cover of silver fir (*Abies amabilis*). Coarse-scale limitations on seed availability, germination and establishment are posed by the prominence of silver fir in the surrounding area and by interannual climate variability. At subalpine elevations, successional processes are similarly drawn out, but here they also vary at fine spatial scales as micro-habitats differ in their suitability for plant establishment in a given year. Succession results in vegetation classification ambiguity because most vegetation follows a gradual trajectory with no clear and repeatable breaks between stages. Third, at high elevations, the spatial grain of available habitat for individual species is so fine that it becomes impractical to delineate all distinct assemblages of

[1] As contrasted with the more discontinuous influence exerted by factors such as soil chemistry.

species. Here, the vegetation might better be described as a variable mosaic in which assemblage dominance tilts across more coarsely scaled gradients. This results in field interpretation challenges because of lack of clarity about the minimum patch homogeneity and size needed to constitute a sampling unit. At these elevations each species in fact exploits micro-habitat niches that become available to it on an individual basis.

The starting point for the NCCN map classification was an early draft of the interim NCCN alliances later presented in NatureServe (2012). These draft interim alliances were defined by a crosswalk from the vegetation associations presented in Crawford et al. (2009). Those associations, in turn, were defined by using *classification plots*—collected from 2005 to 2007 at all NCCN parks—to refine and provide context to several previous regional classifications. Collection of *mapping plots*, which were field-assigned to vegetation association using the keys in Crawford et al. (2009), began in 2008, using draft versions of those keys. The original plan was to train map models based solely on mapping plots, but it quickly became clear that not enough mapping plots had been collected—particularly at MORA—to adequately train models, so the previously collected classification plots were added to the map training pool. The possibility that these plots were assigned to associations based on criteria other than the keys in Crawford et al. (2009) presented a potential downside to their use for this purpose.

Early map modeling results found a significant degree of mismatch between the assigned alliance of many plots and their modeling tendencies. Through experimentation, it became clear that many of the errors were a consequence of dissimilar plots, particularly in conifer forests, being assigned to the same alliance due to key breaks that resulted in artificial boundaries between types. Although the associations were originally derived via multivariate cluster analyses and are generally "bloblike" in *n*-dimensional space, the keys carved straight lines through these concepts, squaring them off with hard breaks such as "*Oplopanax horridus* > 5%." Using key-based calls lowered floristic cohesion within the resulting associations and resulted in many plots that modeled poorly as the class to which they had been assigned. We addressed this problem by moving to a multivariate clustering approach for determining the best floristic match for a given plot.[2]

Field data from all plots collected after 2008 included reasonably complete species cover data, which allowed us to retroactively reassign association calls if the overall species composition warranted that. When this dataset was completed for all three parks, it also enabled us to revise the associations themselves in order to correct a variety of pre-existing issues. We termed the revised concepts *mapping associations*, as we were unsure whether the NVC would be adjusted to incorporate them (many have in fact been included in Ramm-Granberg et al. 2021). Regardless, the revisions were a necessary step in rolling the plot-level data up into a mappable classification. The mapping associations were combined into map classes based on their joint floristic and modeling similarities.

[2] Although we have provided a dichotomous key for field use in identifying the final map classes (in keeping with NPS required deliverables), dropping the use of keys for assigning vegetation associations was an essential step in deriving map classes with floristic and modeling cohesion.

Despite the additional steps involved in their production, many of the mapped classes bear strong resemblances to the original concepts presented in NatureServe (2012).

Another classification challenge we encountered was that areas recently disturbed by fire, flooding or mass movement sometimes fit awkwardly into the original NVC associations. We moved away from strict floristics-based labeling of plots affected by disturbance and considered their setting and site history as well. Several iterations of plot-level examination and reassignment followed by association-level floristic recalculation resulted in convergence of the classification on several floristically consistent associations often connected with natural disturbance. Past anthropogenic disturbance, such as selective logging of Sitka spruce (*Picea sitchensis*) near the coast,[3] also occasionally resulted in ambiguities. Moving away from key-based assignment of association calls helped considerably here: in the case of logged Sitka spruce, enough floristic signals persisted from the natural vegetation community that the correct call was evident, even if the spruce cover was now well below ten percent. Finally, areas experiencing ongoing change (e.g., conifer encroachment into established meadows) also present a challenge as the combination of species may not have been well-represented in the past. Some flexibility is necessary in such areas, which are likely harbingers of greater dilemmas to come.

1.2.2. Mapping

In contrast to many VMI products, we mapped the large NCCN parks using automated model-based methods rather than photo-interpretation (PI). This decision was originally made because of the size of the parks and the indistinct appearance of many of the map classes in imagery. For example, two-thirds of the NCCN parks are covered by coniferous forests and woodlands of 24 different map classes. The component tree species generally cannot be visually distinguished, and many are recognized in the field based as much on their understory composition. In addition, the gradual change of species prominence along climatic gradients in PNW forests, and the variable and patchy species composition characteristic of many non-forest patches, result in a landscape that is not easily divided by hand into discrete patches.[4] Despite this, and the inherent classification challenges discussed above, sites within the parks can be broken repeatably into map classes based on their full species composition, and these classes can be reliably mapped using model-based techniques.

We used an inductive modeling process, in which a computer learns how to distinguish map classes by the examples provided from field plots. We used the *random forests* algorithm (RF; Breiman 2001), as adapted for the R language (Liaw and Wiener 2002). The large number and unequal abundance of map classes proved to be a challenge to multi-class models, which were unable to simultaneously perform well at the prediction of all classes. To address this, we decomposed each multi-class model into many *one versus one* binary models, in which individual map classes were

[3] Because the same classification pertains to MORA, OLYM and NOCA, throughout the text we have chosen examples to illustrate our approach from across the three parks (although it is applicable to OLYM in this case).

[4] Many non-forest classes might have been mapped by hand with ideally timed high-resolution aerial imagery. However, we did not feel that the expense of collecting such imagery over the large expanse of the parks could be justified for an uncertain outcome that still would have left the forested lands unmappable via photo-interpretation.

modeled directly against each other (see Bishop 2006, p. 339). We used a novel predictor selection scheme that reduced prediction time, limited collinearity in the predictive variables, and co-optimized model accuracy and effective spatial resolution. The model-based results were manually edited where needed[5] and then subjected to a map accuracy assessment based on independent field data collected based on a stratified random sample.

1.2.3. Spatial resolution and minimum mapping unit

Several map classes often occur in patches of 100 m^2 or smaller. We attempted to capture these occurrences, and to produce a map resembling manually delineated VMI maps, by modeling on 3-meter pixels (9 m^2). We smoothed the raw model outputs and filtered to a class-specific minimum patch size ranging from 81–441 m^2. Many occurrences above those thresholds likely remained undetected, because some essential predictors were derived from coarser resolution sources. Based on the average resolution of the predictors selected across all models, a typical minimum mapping unit (MMU) of 500 m^2 can be assumed, although many patches smaller than that are mapped.

1.3. Project area

1.3.1. Geography

Olympic National Park (OLYM) occupies much of the Olympic Peninsula of northwestern Washington state, lying about 130 km (80 miles) west of Seattle. It is the largest of the North Coast and Cascades Network parks—with which it is shown in **Figure 1**—and the sixth largest national park in the contiguous United States. It is adjacent to other protected lands, including several small wilderness areas in the Olympic National Forest, but is also bordered by extensive private and public lands managed primarily for timber production.

[5] Polygons of some distinct yet rare vegetation types (e.g., ruderal meadows at old farm sites at OLYM) modeled poorly due to insufficient training data, but were easily reassigned by hand. Strips adjacent to roads also frequently mapped poorly and were reassigned manually.

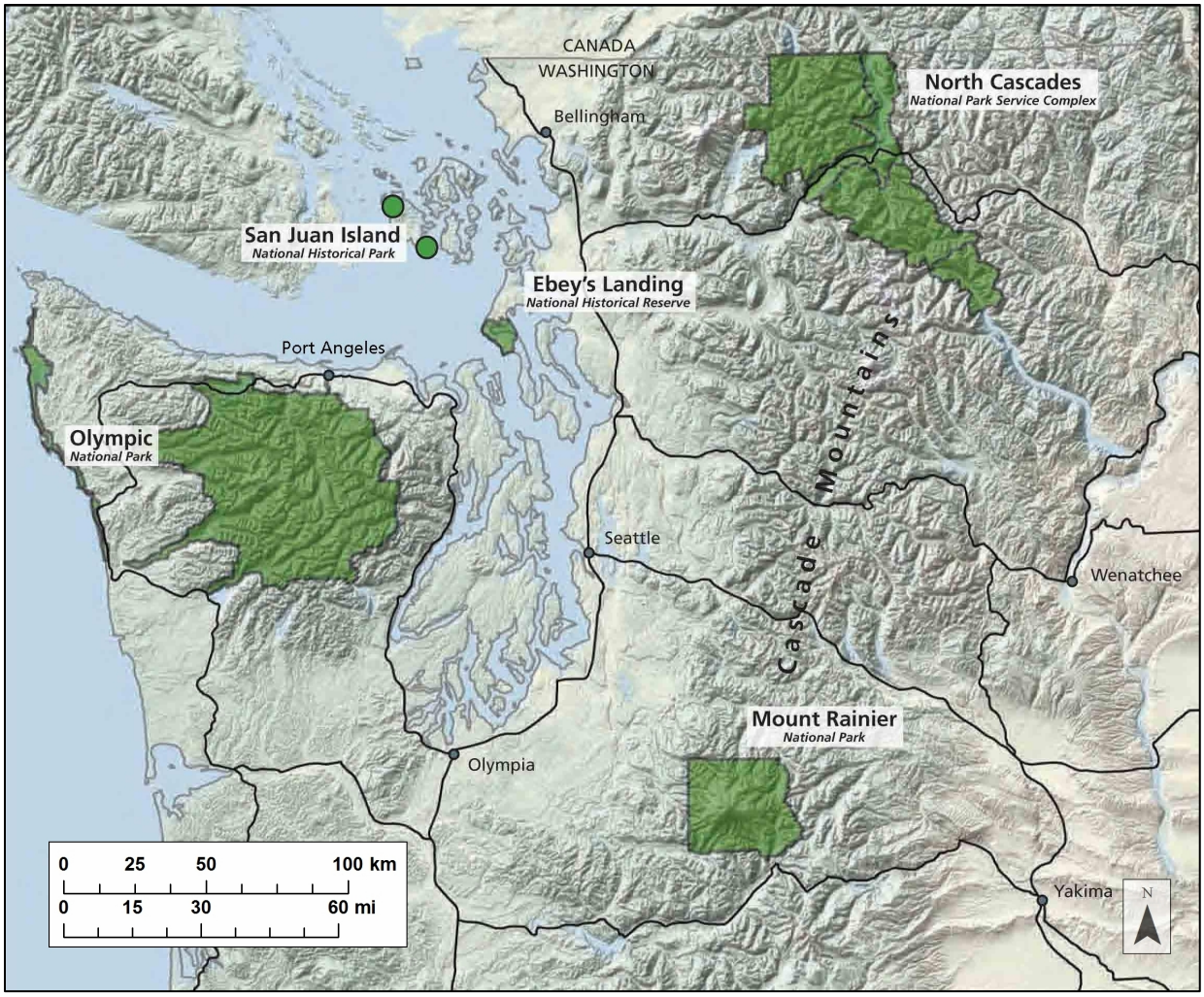


Figure 1. Map of North Coast and Cascades Network National Parks from Copass and Ramm-Granberg (2016b). Fort Vancouver National Historic Site and Lewis & Clark National Historical Park lie farther south and are not shown.

The park (**Figure 2**) combines a narrow coastal strip of wilderness along the Pacific Ocean, lush lowland river valleys with temperate rainforest, and a mountainous interior with extensive subalpine and alpine habitats. Ten major rivers radiate from the central core of the park, which is anchored by Mount Olympus (2,428 meters; 7,966 feet). The Elwha River flows north through the park interior for 58 kilometers (36 miles), collecting water from many of the highest peaks. The Bogachiel, Hoh, Queets and Quinault drain the west side, where a cool and moist maritime climate extends far up the valleys, supporting temperate rainforest and marked by frequent coastal fog. The Dungeness, Dosewallips, Duckabush and Skokomish flow east, and the Wynoochee runs south. The northeast section of the park, east of the Elwha, lies in an intense rain shadow cast by the mountains in the central core. Conditions are much drier and vegetation communities are quite distinct from the rest of the park, bearing some resemblance to those of the eastern Cascades. Elevations in the park range from sea level to the summit of Mount Olympus.

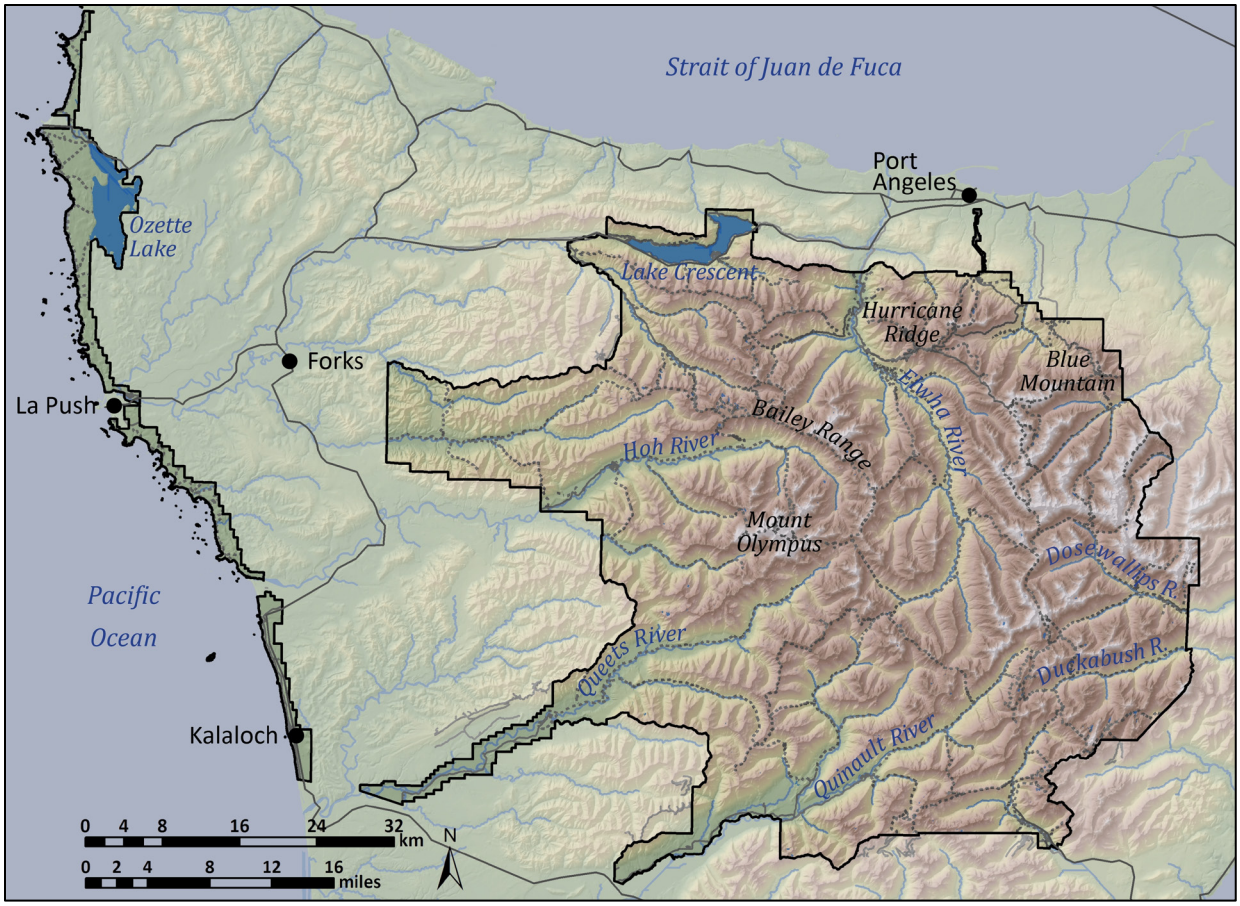


Figure 2. Map of Olympic National Park, illustrating topography, rivers, roads and other features.

We defined the project area as the park (369,901 hectares; 914,405 acres), in addition to a variable-width buffer determined by availability of key predictor geospatial data. The surrounding buffer area, which was not assessed for map accuracy, accounts for 39% of the total project extent of 606,147 hectares (1,497,822 acres).

1.3.2. Environmental setting, bioclimatic zones and major vegetation types

The Olympic Mountains were formed 20–30 million years ago, when the Juan de Fuca plate collided with the western edge of the North American continental plate (Williams 2002). As the oceanic plate subducted under the continental plate, basalt and sedimentary rock layers on the seabed floor were scraped off, resulting in a distinct horseshoe shape with a core of metamorphosed sedimentary rocks ringed by basalt. Later, repeated glaciation during the Pleistocene scoured the rocks and shaped the peaks as seen today. The final continental ice advance wrapped around the northern and eastern parts of the range, damming many rivers and creating large lakes in the river valleys, while leaving the northeastern peaks isolated. These have served as glacial refugia for species that persist to current times, contributing to a high rate of endemism (Buckingham et al. 1995).

The park's west side receives 300–500 cm (120–200 in) of annual precipitation, with the maximum precipitation recorded at Blue Glacier on Mount Olympus. In contrast, the east side can receive as

little as 100 cm (40 in) (Henderson et al. 1989). Below 300 meters elevation, winter precipitation falls exclusively as rain. Higher, up to 750 meters, it falls as mixed rain and snow, while above that it is primarily snow. Temperatures are moderated by proximity to the ocean. Mean January minimum temperatures in the coastal strip are above freezing, ranging up to 4 °C (39 °F), but are as low as -8 °C (18 °F) at high elevations on the east side. Average July maximums are from 16–20 °C (61–68 °F) at the coast, but up to 25 °C (77 °F) in low elevation valleys in the park interior (Baccus 2018, Davey et al. 2006, PRISM Climate Group 2019).

Topography, climate and glacial history drive significant plant diversity, with 1,119 native vascular taxa documented on the peninsula (Buckingham et al. 1995). Because the park ranges from the coast to the alpine, it certainly includes most of these. The varied landforms, from coastal bluffs and beaches to glacial peaks, in combination with steep temperature and precipitation gradients, result in varied and complex vegetation patterns (Henderson et al. 1989). The major vegetation zones are discussed below.

Lowland forests

Along the coast, forests of Sitka spruce (*Picea sitchensis*), western hemlock (*Tsuga heterophylla*) and western redcedar (*Thuja plicata*) are found on gentle slopes. These frequently have an impenetrably dense understory of evergreen huckleberry (*Vaccinium ovatum*) and salal (*Gaultheria shallon*), with pockets of skunk-cabbage (*Lysichiton americanus*) and slough sedge (*Carex obnupta*) in wet spots. Moving inland up west-side drainages, a coastal fog belt sustains forests of Sitka spruce with western hemlock, western redcedar and Douglas-fir (*Pseudotsuga menziesii*). These forest canopies have complex structure and the trees can be massive. Above or inland from the coastal fog belt, Douglas-fir and western hemlock-dominated forests range from about 300 to 600 meters elevation; their understories vary depending on site moisture and stand history. On the east side, lowland forests range to significantly higher elevations. Broadleaf riparian forests dominated by red alder (*Alnus rubra*) and black cottonwood (*Populus trichocarpa*) are found in floodplains adjacent to these conifer stands, while bigleaf maple (*Acer macrophyllum*) is often prominent in upland areas recovering from disturbance.

Lower montane forests

The lower montane zone is defined primarily based on the codominance of silver fir (*Abies amabilis*) with western hemlock and the rarity of tree species associated with higher elevations. The lowest occurrences of silver fir are in moist forests on valley bottoms and north-facing slopes. The bulk of the zone, however, occurs in valley wall settings. Mesic forests characterized by silver fir and western hemlock predominate up to about 1,100 meters elevation. This is the most widespread zone in the park, occupying vast valley wall areas on the west side. Toward the east, however, it is increasingly constricted as lowland forests extend higher. It is essentially eliminated in the northeastern section of the park, where lowland forests contact the upper montane zone.

Upper montane forests

The upper montane zone is characterized by closed forests with substantial cover of higher elevation species such as mountain hemlock (*Tsuga mertensiana*), subalpine fir (*Abies lasiocarpa*) and Alaska-cedar (*Callitropsis nootkatensis*). These forests are predominant up to about 1,350 meters. Tall

shrublands of Sitka alder (*Alnus viridis*), often with Alaska-cedar, are frequent in avalanche tracks and other disturbed areas. Rich herbaceous meadows with cow parsnip (*Heracleum maximum*), edible thistle (*Cirsium edule*) and western meadowrue (*Thalictrum occidentale*), occur on toe slopes in some drainages in this elevation range. In the northeastern section, the upper montane zone is entered directly from the lowland forest zone, as increasing amounts of subalpine fir join the Douglas-fir that characterizes lower elevations there.

Subalpine environments

Above the closed forest zone, upper montane forests transition into subalpine woodlands and tree islands. Subalpine *parklands*, mosaics of wooded areas with dwarf shrublands and meadows, are predominant from about 1,350 to 1,600 meters elevation. Dominant trees are mountain hemlock and Alaska-cedar on the west side, and subalpine fir on the east side. On the driest east-side ridges, whitebark pine (*Pinus albicaulis*) joins subalpine fir. Subalpine shrublands are typically dominated by dwarf ericaceous shrubs such as pink mountain-heather (*Phyllodoce empetriformis*), Cascade blueberry (*Vaccinium deliciosum*) and white mountain-heather (*Cassiope mertensiana*), with some taller shrubs such as big huckleberry (*Vaccinium membranaceum*) and Sitka mountain-ash (*Sorbus sitchensis*) in protected areas. Herbaceous subalpine vegetation is represented by several meadow types, with common species including showy sedge (*Carex spectabilis*), subalpine lupine (*Lupinus latifolius*), American bistort (*Bistorta bistortoides*), Sitka valerian (*Valeriana sitchensis*), wandering daisy (*Erigeron glacialis, E. peregrinus*), Parry's rush (*Juncus parryi*) and Olympic aster (*Eucephalus paucicapitatus*).

Alpine environments

In the alpine zone, which generally ranges from about 1,600 meters elevation upward, tree cover is reduced to stunted krummholz of mountain hemlock, subalpine fir, and Alaska-cedar. Dwarf shrublands and meadows transition into sparser alpine variants with shorter vegetation and fewer species. Common species include spreading phlox (*Phlox diffusa*), yarrow (*Achillea millefolium*), common juniper (*Juniperus communis*), partridgefoot (*Luetkea pectinata*), Tolmie's saxifrage (*Micranthes tolmiei*) and Roemer's fescue (*Festuca roemeri*). Eventually vegetation gives way almost completely to barren bedrock, talus, permanent snowfields and glaciers.

1.3.3. Human history

Evidence of early Native American presence on the peninsula dates back at least 12,000 years. The landscapes that are now included in the park are linked intrinsically to the cultural identity of the present-day Makah, Quileute, Hoh, Quinault, Skokomish, Port Gamble S'Klallam, Jamestown S'Klallam, and Lower Elwha Klallam tribes (NPS 2019). After the end of the most recent glaciation, approximately 5,000 years before present, regional climate and sea level stabilized and modern vegetation patterns emerged. Bountiful maritime resources, regular salmon runs and extensive food resources in the prairies and forests supported a thriving indigenous population. There is extensive documentation of the historic use of many plants as well as modification of the landscape using fire, including in coastal bogs (Reagan 1909, Norton 1979, Anderson 2009).

Euro-American settlement of the peninsula began in the mid-1800s and the remarkable forests attracted early logging around the outer ring of the peninsula. Conservation of the mountains and

forests at the heart of the Olympic Peninsula began with the creation of the Olympic Forest Reserve in 1897, the center of which was designated Mount Olympus National Monument in 1909. As concerns about over-harvesting of elk grew, Olympic National Park was founded in 1938 "to preserve...the finest sample of primeval forests of Sitka spruce, western hemlock, Douglas-fir, and western redcedar in the entire United States" (U.S. Senate 1938). The wilderness coastal strip and a portion of the Queets River watershed were added to the park in 1953.

1.3.4. Previous vegetation studies

Botanists accompanied the early Euro-American explorations of the Pacific Northwest and made collections on the fringes of the Olympic Peninsula. Louis Henderson, with the 1890 O'Neil Expedition, was the first botanist to survey within current park boundaries. Continuing botanical surveys culminated in the publication of the first flora of Washington (Piper 1906). The earliest detailed accounts of the landscape, forest composition and structure, and fire history of the park are found in the survey reports of Dodwell and Rixon (1902). Though not specific to the park, Buckingham (1995) provides an updated account of the flora of the peninsula.[6]

Vegetation classification work began with an account of vegetation community transitions (Murie 1935) and a description of vegetation patterns based on the Merriam life zones concept (Jones 1947). Fonda and Bliss (1969) published a classification for Olympic montane and subalpine forests, and Smith and Henderson (1986) created a lowland forest classification based on inventories in the Hoh and Dosewallips watersheds. Henderson et al. (1989) followed this with a general classification for Olympic Peninsula forests. Kuramoto and Bliss (1970) classified subalpine meadows; Houston et al. (1994) expanded on this work in describing the impacts of non-native goats on subalpine communities. As part of her regional work on coastal bogs and wetlands, Kunze (1994) classified many large wetlands within the park. Many of these classification results, as well as other regional studies, were incorporated into the vegetation classification portion of this project (Crawford et al. 2009, updated by Ramm-Granberg et al. 2021). Other recent publications that are relevant to our mapping work include a detailed examination of the processes involved in riparian forest development along the Queets River (Van Pelt et al. 2006), and a review of the fire history of the Elwha River watershed and of the park more broadly (Wendel and Zabowski 2010).

Vegetation mapping efforts have proceeded in fits and starts over the last 120 years. Dodwell and Rixon (1902) included maps that were digitized by OLYM staff in 1997. Later, detailed maps of dominant forest canopy species and disturbance impacts were produced across the entire Douglas-fir zone in the Pacific Northwest (Andrews and Cowlin 1936). The next spate of mapping came with the increasing availability and ease of use of satellite imagery in the 1980s. Cibula (1981) mapped the park using 60-meter resolution Landsat MSS data, later expanding the effort to incorporate topographic and climatological modeling (Cibula 1987). This work resulted in a total of 21 map classes based on nine distinct spectral signatures. Pacific Meridian Resources used a modified supervised classification method to map 15 vegetated and three unvegetated classes, in addition to some elements of vegetation structure, from 30-meter resolution Landsat TM imagery collected in

[6] Buckingham (1995) lists 1,452 vascular plant taxa on the peninsula; of these 1,119 are native.

1991 (PMR 1997). The classification focused on forests; all herbaceous and shrub-dominated vegetation were lumped into a single class each. Finally, Henderson et al. (2011) used a range of environmental variables to predict boundaries between eight potential natural vegetation (PNV) zones on the Olympic Peninsula. Evaluating the outcomes of the current project in the context of Henderson's work would be a worthwhile endeavor.

1.4. Project timeline

The following timeline describes the primary activities during each year of the 15-year project. Only activities at OLYM are described; activities were focused on other NCCN parks during several years.

2005 — NPS project initiation, planning and scoping, fieldwork for accuracy assessment of previous generation vegetation map (PMR 1997), database development.
2006 — Classification fieldwork, planning and scoping, database development.
2007 — Classification fieldwork, development of association-level NVC and database.
2008 — INR joins project. NAIP and satellite image collection and processing, development of mapping fieldwork protocols, development of association-level and higher-level NVC.
2009 — Development of map training fieldwork protocols, limited map training fieldwork, development of higher level NVC, satellite image collection and processing, development of predictor metrics methodologies.
2010 — Map training sample design and fieldwork, NAIP image collection and processing, development of predictor metrics methodologies.
2011 — Map training fieldwork, development of predictor metrics methodologies.
2012 — Map training fieldwork for remaining poorly sampled areas, training data quality control, NAIP and satellite image collection and processing.
2013 — Production of draft vegetation map for use in stratification of accuracy assessment sampling, AA sample design, AA fieldwork, training data quality control.
2014 — Training data quality control, AA data quality control, NAIP image collection and processing.
2015 — NPS gives INR go-ahead to approach NCCN projects as a single entity and to work on the classification as needed for successful mapping. Training data quality control, AA data quality control, floristics methods development.
2016 — Training data quality control, floristics quality control, mapping associations development, NAIP image collection and processing.
2017 — Training data quality control, floristics quality control, mapping associations development, map classification development, development of new topographic predictor metrics, refinement of nested texture metrics methodology, lidar data acquisition and processing.
2018 — Training data quality control, mapping associations development, map classification development, satellite image collection and processing, draft map production.
2019 — NPS draft map review. Training data quality control, mapping associations completion, map classification completion, AA data quality control, development of shadow correction methods for NAIP imagery.
2020 — Production of final maps, AA analysis, report.

2. Methods and Results

Most NPS VMI maps have been produced by photo-interpretation (PI). We used model-based methods instead, because of the large size of the NCCN parks and the visual similarity of many of the key plant communities. Machine learning methods were used to extrapolate from a large set of classified field plots to the full extent of the park. The mapped vegetation units were 3-meter pixels rather than polygons, because we found that pixel-based modeling was the only reliable method for boundary detection between visually similar map classes. We invented and developed a variety of innovative image processing and modeling techniques to achieve finer spatial resolution and greater accuracy than is typical of model-based vegetation maps. The primary phases of the mapping process—many of which occurred concurrently—were collection and basic quality control of training field data (**Section 2.1**), floristics data treatment and associated plot QC (**Section 2.2**), development of mapping associations and associated plot QC (**Section 2.3**), crosswalking associations to map classes (**Section 2.4**), acquisition and pre-processing of predictive data sources (**Section 2.5**), development and creation of predictive metrics (**Section 2.6**), machine learning-based modeling (**Section 2.7**) and post-processing and editing (**Section 2.8**).

2.1. Field data

Field data were used to develop an association classification (**Section 2.3**) and map classification (**Section 2.4**) and to provide training data for the machine learning processes used to create the map (**Section 2.7.2**). The same classification was used at each of the large NCCN parks; its development drew from 4,110 plots collected at MORA, OLYM and NOCA. Because the classification relied on data collected from all parks, each of those protocols is reviewed here. The map training data included 2,007 field plots, all collected at OLYM, in addition to the PI plots discussed later.

2.1.1. Sample collection

We trained predictive models using plots from multiple sampling efforts with distinct sample designs and field protocols. Most were collected during the mapping phase of the project, between 2008 and 2015. Although field protocols evolved over this time, the fundamentals were in place by 2010.[7] We also used many plots collected in 2006–07, during the initial NVC development phase of the project (see Crawford et al. 2009). Many of these plots included full floristic data and suited our needs well.

Training data for inadequately sampled vegetation types were supplemented by incorporating plots from a variety of other field efforts in the parks between 2005 and 2015. Although the protocols varied widely for these plots, through the quality control process (**Section 2.1.2**) we converted all data to a standardized format: a circle of known radius georeferenced to aerial imagery collected in 2015 (**Section 2.5.1**) and a species list with cover estimates to the nearest one percent (or *trace* if

[7] The primary requirements were (a) plot dimensions adjusted to match a homogeneous vegetation patch, up to a maximum 40-m radius circle; (b) documented plot center location and radius in four cardinal directions; (c) diagram illustrating landmarks and land cover transitions, for spatial QC; (d) reasonably complete floristic data including visual percent cover estimates; and (e) photos at cardinal directions from center, for spatial and floristic QC.

present in smaller amounts). Plots with reasonably complete species occurrence data were designated *full-ocular* plots, while those with incomplete data were designated as *partial-ocular*.

The following sections outline the sample designs, field protocols, and data QC procedures for all plots used to create the map classification or the OLYM vegetation map. **Table 2** summarizes this information. The table does not include photo-interpreted plots, which were mainly used as training for unvegetated classes.

Table 2. Total number of field plots used to create the map classification ("full floristics plots") and the OLYM vegetation map ("model plots"), categorized by park and collection effort. "Ocular type" specifies whether documenting full species cover data was an objective of the protocol. Photo-interpreted plots were also used for modeling; those are not included here.

Collection effort	Collection years	Collected by	Ocular type	Full floristics plots			Model plots [A]
				MORA	OLYM	NOCA	OLYM
VIP classification	2006–07	NPS	Full	186	228	79	246
VIP mapping protocol X	2008	NPS	Partial [D]	44	0	0	0
VIP mapping protocol M [B]	2009–11	NPS	Partial [D]	33	10	0	69
VIP mapping protocol Y	2009–11	NPS	Full	151	1,094	0	1,199
VIP mapping protocol Z [C]	2012–15	NPS	Full	0	233	1,612	255
VIP mapping protocol Q	2014, 2019	INR, NPS	Partial [D]	1	0	0	3
PMR accuracy assessment	2005–06	NPS	Full	61	18	10	135
Monitoring reconnaissance	2005–14	NPS	Full	91	46	48	100
UW forest community	2015	UW	Full	165	0	0	0
Total	**–**	**–**	**–**	**732**	**1,629**	**1,749**	**2,007**

[A] There is a large amount of overlap with full floristics plots collected at OLYM. Totals include additional training data created in adjacent or included patches based on plot notes.

[B] Totals reflect number of individual patches from subdivision of original mosaic plots.

[C] Including *verification plots*, revisits and updates to previously collected plots using revised protocols.

[D] Some plots had floristics supplemented later by inspection of field photos, and were treated as full-ocular plots.

NCCN VIP plot types

The following plot types were collected especially for the NCCN VIP, and are listed chronologically.

Classification (2006–07)

These plots were intended primarily to support development of the NVC for the NCCN (Crawford et al. 2009, Ramm-Granberg et al. 2021) and were collected by NPS at the three major parks. Crews sampled a broad range of types.[8] The protocol included collection of a comprehensive species list

[8] Data collection was particularly focused on vegetation types known to be undersampled in the existing draft classification, such as shrub-dominated avalanche chutes. Circular plots were located opportunistically in homogenous patches that were large enough to meet plot size recommendations. Forested plots were sampled over

with cover estimates and several field photos. Plots were assigned to an association from an early draft of the NVC, or to a provisional association if no good match could be determined.

Mapping protocol X (2008)

Collected only at MORA, these plots represented the earliest sampling implemented during the VIP mapping phase, before significant improvements were made in spring 2009. Polygons segmented from true color aerial imagery were targeted for opportunistic sampling. Estimates of crown cover were made for up to three tree species, but understory species were generally not documented. At least two photos were taken. A vegetation association was selected from an early draft of the NVC.[9]

Mapping protocol Y (2009–11)

A variety of changes to the sample design and field protocols were implemented beginning in the 2009 field season. A stratified sample design was implemented to guide the effort.[10] Map sheets produced from resolution-merged aerial and satellite imagery allowed crews to navigate more efficiently, locate plots more accurately, and document vegetation patches for later use; similar paper maps were used in all subsequent field efforts (see **Figure 3** for an example). Beginning with protocol Y, plot dimensions were determined by the extent of the homogeneous vegetation patch present at plot center, up to a maximum 40-meter radius circle. Vegetation transitions along four perpendicular transects from the plot center were documented, and the plots were drawn by hand on the map sheet and documented in greater detail in a field diagram on the data sheet. Species cover was visually estimated for most plant species present.[11] Photos were taken at cardinal directions from plot center, and the best-fit vegetation association was selected from the newly published NCCN NVC classification (Crawford et al. 2009). Protocol Y was the main source of map training data at OLYM; a completed fieldsheet is shown in **Figure 4**.

---

an 11.3-meter radius (400 m^2), woodlands and shrublands over an 8.0-meter radius (200 m^2), and herbaceous and sparsely vegetated plots over a 5.6-meter radius (100 m^2). Notes on soil conditions and fire history were taken.

[9] A plot center was selected within a representative homogeneous area for assessment of the vegetation association and canopy composition, which was documented within a 20-meter radius circle around the point. A secondary association was listed if there was a clearly distinct association located nearby, but no corresponding location information was documented.

[10] Sample sites were targeted by using an unsupervised classification technique to break parks into fifty distinct strata based on Landsat reflectance data, topographic metrics, and geographic blocks. Within each stratum, targets were established in the most homogeneous Landsat pixel clumps within several hundred meters of trails. Field crews navigated to these locations and also established opportunistic plots in homogeneous occurrences of vegetation types that had been poorly represented in the targeted sampling.

[11] Over the years, field crews were increasingly comprised of returning, experienced members, and the capability to collect complete species composition data increased. The OLYM protocol Y data (collected primarily in 2010-11) therefore have considerably greater completeness than that from MORA (collected primarily in 2009). NPS's original aim had been to produce a completed map of MORA before the other parks, so the MORA fieldwork was done in "hurry-up" mode, and a return there with the more experienced crews was never realized.

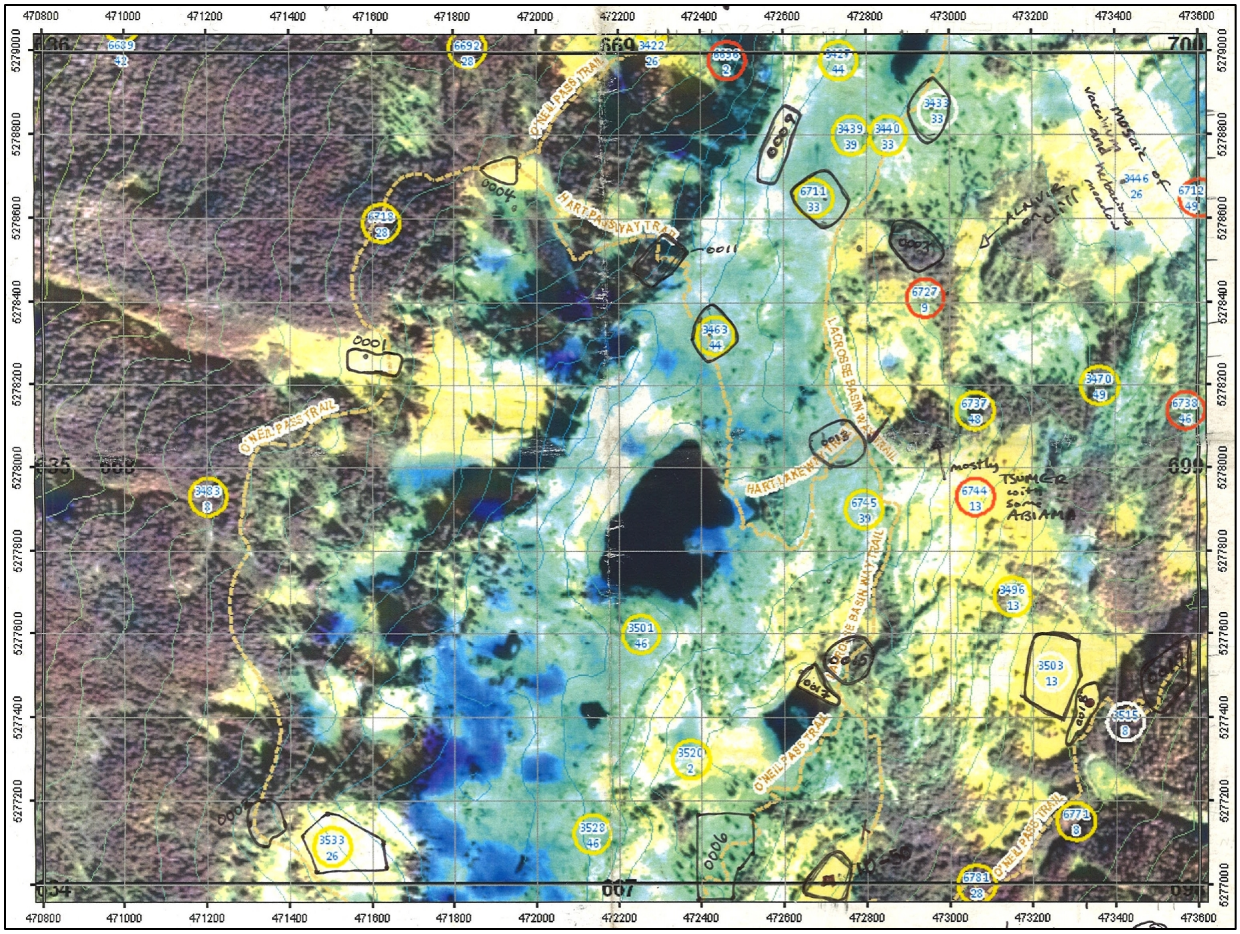


Figure 3. Completed mapping phase map sheet(protocol Y, 2011).

2009 NCCN Veg Mapping Project — version: 8/15/2010

Park: OLYM	Map Number: 395	Plot # 0002	Cluster OPP
Date: 6-25-2011	Observers: ML		

GPS Taken (Y) N	GPS notes		
GPS Unit: A	UTM E 448158	UTM N 5300187	Coordinate source: (GPS) / Map
Plot Type:	NAV / (OPP)	Obs Type: (Onsite) / distance	Plot Size: (Large) / narrow / small

Micro Position					
____ toe of slope	____ bench/flat	____ ridge top	____ river bottom	X middle 1/3	____ draw/stream bottom	____ upper 1/3	____ lower 1/3	____ wetland

Slope: 28 deg	Aspect: 296°

Polygon Added/Modifed: (Y) / N	
Imagery Notes:	

Plot Transect	DISTANCE (m)	CHANGE (IF ANY)
NORTH	200	
EAST	80	
SOUTH	180	
WEST	140	

Site Description: Closed canopy multiple age class stand codom. by ABIAMA + TSUHET. Some TSUMER + CHANOO present in understory. ABIAMA dominates regen layer. Shrub layer poorly developed mostly Vaccinium sp. Herb layer just starting to leaf out due to late snow, so cover could be higher. TSUHET makes up largest sized cohort (>60 dBH) w/ ABIAMA (20-50 dBH) in the canopy.

Inclusion Description:

Disturbance Description: → Small shrubby ALNVIR + RUBSPE patches where canopy openings occur.

Photos ? (Y) / N	Direction	Image #	Description
	NORTH	2897	green patch (R) side
	EAST	2898	broken small snag (R)
	SOUTH	2899	looking upslope
	WEST	2900	looking downslope

Fuel Type: TL3	Fuels Notes: light load of 1-3" wood.

Group code: Z41	Alliance Code: 007	Association:
Key worked: (Y) / N	Key Notes:	
Description good: (Y) / N	Description Notes:	

Trees		Cover
T1	ABIAMA	40
T2	TSUHET	30
T3		
T4		
T5		
T6		
Total overstory tree cover		70

Trees		Cover
U1	CHANOO	2
U2	TSUHET	8
U3	ABIAMA	15
U4	[illegible]	[illegible]
Total understory tree cover		25

AVG Overstory Tree DBH (cm)	40

Shrubs		Cover
S1	VACALA	5
S2	VACPAR	5
S3	RUBPED	5
S4	RUBSPE	2
S5		
S6		
Total shrub cover		17

Herbaceous		Cover
H1	TIATRI	3
H2	STRROS	7
H3	CLIUNI	3
H4	ACHTRI	3
H5	MAIDIL	2
H6	POLMUN	2
H7	VERVIR	3
H8	Moss sp.	15
Total herbeceous cover		38

Abiotic (if >10%)	Cover
Snow/Ice	
Bare Soil	
Rock	
Other	
Total Abiotic cover total	

QA:

☑ Data Sheet entered ☐ Data Sheet verified ☐ photos verified ☐ plot digitized ☐ scanned

Figure 4. Completed mapping phase data sheet(protocol Y, 2011).

Mapping protocol M (2009–11)

Field crews targeted subalpine and alpine areas to determine whether fine-scale mosaics of distinct alliance level vegetation might be combined into recurring mappable types.[12] This effort targeted mosaics of vegetation patches, each of which was smaller than those considered for sampling during prior efforts. Species cover was estimated for the most significant species in each distinct patch and an association was chosen from Crawford et al. (2009). The data were later analyzed to assess whether the patches might be combined into consistent coarser-scale vegetation types, but patterns were not consistent enough to allow this.

Mapping protocol Z (2012–14)

Field sampling was guided by a revised stratified sample design.[13] The field protocol was similar to protocol Y, except that association transitions along each transect were documented in great detail. This allowed additional plots to be generated later if needed. Field crews were remarkably stable during this period, allowing the collection of more complete species cover data at nearly all plots. The data collected during this time period was integral to the refinement of the mapping associations (**Section 2.3**).

Mapping protocol Q (2014)

The primary aim of this brief sampling effort was to collect data in accessible but unsampled regions of MORA, which had remained the most poorly sampled park. Early draft maps had difficulties separating forest types containing *Tsuga mertensiana* from those lacking it. To address this problem, we created a species distribution model for *T. mertensiana* and used it to target locations with an intermediate likelihood of presence. We also targeted sites exhibiting a high degree of draft map class uncertainty. Plots were 11-meter radius circles. All species with significant presence were documented, but cover was only estimated for tree species. Understory plants were simply listed in descending order of prominence.

Other plot types

We used plots collected for several other projects to provide data for vegetation types that would have otherwise been inadequately sampled. During the quality control process (**Section 2.1.2**) we adapted the available information to our purposes, making use of field notes, photos, and imagery.

PMR accuracy assessment plots (2005–06)

These plots were collected to assess the accuracy of the previous generation of NCCN vegetation maps (Pacific Meridian Resources 1997). Sampling locations were stratified across the mapped classes. Plots were 28.5-meter radius circles, and were labeled with an association from an early

[12] Early in the project, we had planned to map to the default NPS vegetation inventory MMU of one half-hectare (Lea and Curtis 2010). Most subalpine alliance-level vegetation occurs at considerably finer scales than this.

[13] Combined unsupervised-supervised classification was used to break the landscape into 52 unique strata of Landsat spectral reflectance, climate metrics, and topographic curvature. Again, the most homogeneous accessible areas were determined using an automated procedure, and these locations were targeted on an as-needed basis by crews, in addition to sampling opportunistically in homogeneous vegetation encountered along the routes.

NVC draft or with a provisional call. Cover of the top three species in various height strata was collected.

Monitoring reconnaissance plots (2005–15)

Plots were collected in forests and subalpine areas to assess the suitability of randomly selected locations for long-term monitoring plots. Forest plots were 50x50-meter squares; subalpine plots varied in size. Cover was estimated for dominant species and the surrounding area was coded to an association from the most recent available NCCN classification.

Forest legacy plots (2015)

Plots collected at MORA in the 1970s and 1980s (Franklin et al. 1988) were revisited in a project of Dr. Hille Ris Lambers at the University of Washington. The cover of understory vegetation was estimated over several small subplots, but tree species were documented by counting the number of stems in distinct bole diameter classes rather than by cover.[14] The field notes allowed us to convert these estimates into cover estimates that were reasonably compatible with other plots, and the fairly complete ocular data collected at these plots were critical in providing reference floristics data at MORA, which had been generally undersampled in this regard in earlier efforts. No vegetation type was assigned in the field.

Photo-interpreted plots (2014–19)

We supplemented the field-collected data for several structurally-defined and abiotic map classes[15] by assigning PI locations where it was possible to do so confidently. Generally, we approached this as an iterative process, using previous map generations to assign additional data in areas that appeared to map poorly. We also used targeted absence plots to improve mapping in areas where we could only narrow down the correct answer to one of several map classes but were confident that draft maps were in error. These plots were used only in the binary models (**Section 2.7.3**) which pitted the specified possibly correct classes against the clearly incorrect classes.[16]

---

[14] Understory plants were assessed via a cover estimate in four 1x1-meter quadrats and presence/absence in a 4-meter radius circle; these were converted into an average percent cover for each species. Trees taller than breast height were individually measured in a 12.6-meter radius circle and were summarized in m^2/ha. We converted the stem counts to rough percent cover estimates assuming that crown area was proportional to the square of bole diameter at the individual tree level.

[15] The map classes that received PI plots at OLYM were C26–CONIFER KRUMMHOLZ AND TREED CLIFF, B30–GRAVEL BAR SUCCESSIONAL SHRUBLAND, H58–VEGETATED BALD, R70–COASTAL BARREN, R71–ALLUVIAL BARREN AND DEBRIS-COVERED ICE, R72–COLLUVIAL BARREN, R73–BEDROCK BARREN, W80–SALT WATER, W81–FRESH WATER and W82–EXPOSED SNOW AND ICE.

[16] Examples included meadows on deeply shaded north-facing slopes that had a tendency to map as C26–CONIFER KRUMMHOLZ AND TREED CLIFF, and conifer forests in very flat areas (typically C04–MOIST WESTERN HEMLOCK, DOUGLAS-FIR AND FOAMFLOWER FOREST and C10–MOIST SILVER FIR, WESTERN HEMLOCK AND FOAMFLOWER FOREST) that had previously mapped as W81–FRESH WATER.

The efforts above resulted in the collection of about 6,500 field plots across the three parks. The OLYM field plots used for map training are shown along with photo-interpreted plots in **Figure 5**.

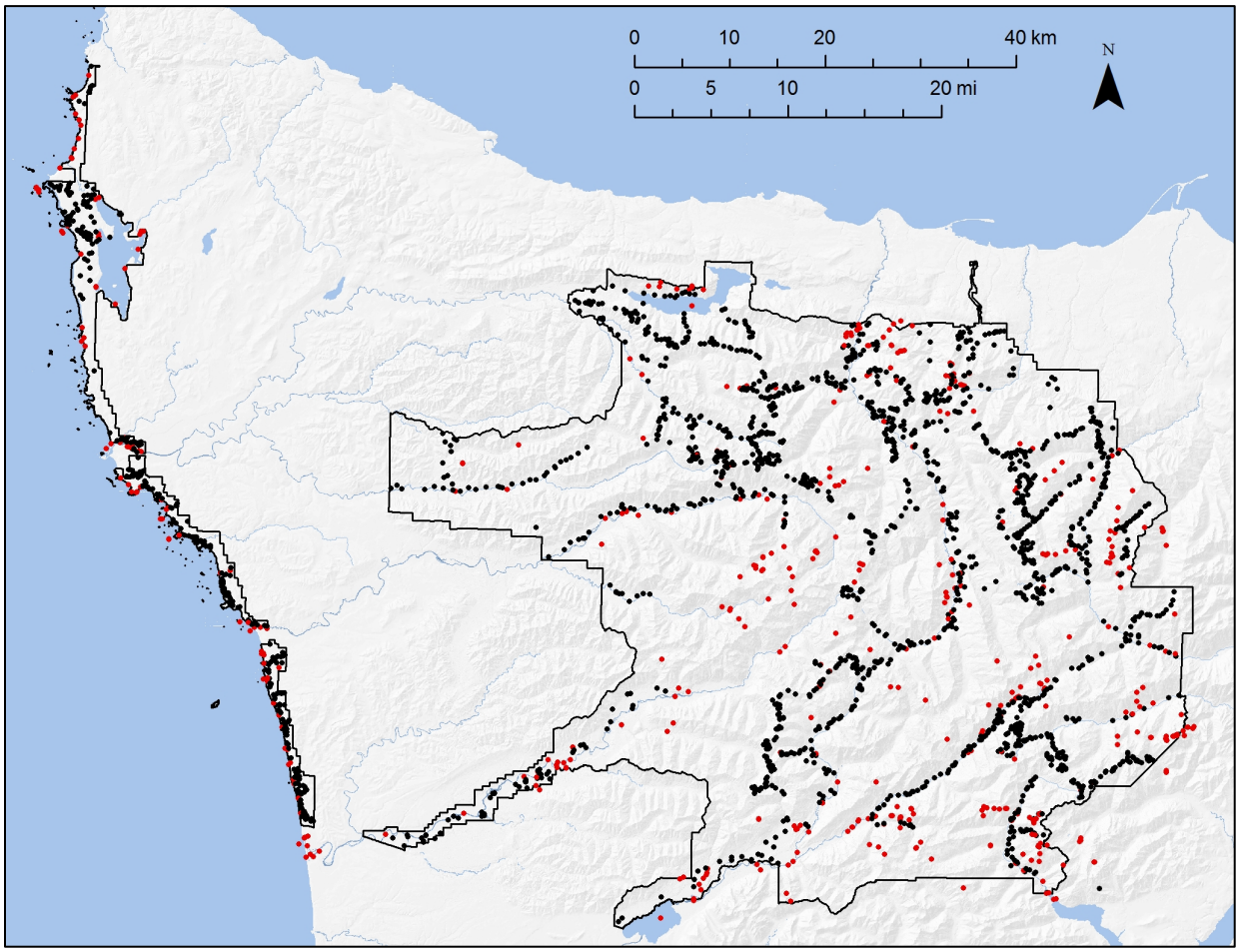


Figure 5. Model training plot locations Of the 2,519 plots, 2,007 were collected in the field (shown in black), and 512 were photo-interpreted (shown in red).

2.1.2. Basic quality control

An extensive quality control process was necessary due to the many distinct field protocols used, the variable field effort applied at different plots (particularly regarding species ocular estimates), ambiguities in patch delineation and ocular estimates resulting from heterogeneous vegetation, spatial inaccuracies due to poor GPS reception, field call misassignments caused by key artifacts, updates to the NVC vegetation classification during the data collection process, occasional species misidentification and data entry errors. The basic quality control steps for training data plots are described below. Quality control of floristics data and vegetation association calls are discussed in **Section 2.2** and **Section 2.3**, respectively.

Spatial characteristics

Because of the fine-scale heterogeneity associated with many vegetation types in the park, we aimed to precisely and accurately locate each training plot with respect to the 2015 NAIP and lidar imagery,

which were the finest-resolution predictor datasets available. Minimizing spatial error was particularly important for non-forest plots in small patches.

The vegetation patches represented by field plots varied in size and shape. In order to simplify data management and modeling, we converted all plots to a circle throughout which the assigned call was applicable, excepting permitted inclusions of 81 m^2 or less (nine 3x3-meter pixels). This allowed us to represent plots simply as a center point and radius. We used a script to identify the center point and radius of the largest circle that would fit within any delineated field polygons. Initial locations and radii were set for other plots based on the protocol's assessment dimensions. For all plots, we verified the spatial characteristics by comparing field notes, plot diagrams, and field photos with NAIP, lidar imagery, and coarser-resolution satellite data.

Plots were flagged for additional review if the GPS center point taken in the field was more than 20 meters from the center point of the field-drawn polygon. We prioritized positioning the circle in the section of the plot closest to the GPS point, assuming that the area nearby was the most thoroughly surveyed portion of the plot. If there were signs of inconsistency between the GPS point, the plot description, and the appearance of the surroundings in imagery, we prioritized the plot diagram and field photos (if provided), repositioning the circle on the plot center as determined by that information.

Disturbance review

Various disturbances impacted the park in the years between field data collection and acquisition of the imagery used for final map production. To prevent the use of training data for which the field-assigned vegetation type no longer corresponded to a plot's condition in imagery, we identified and excluded plots that were disturbed between their sampling date and the acquisition date of the most recent imagery source used in modeling, August 4, 2017.[17] Fires were the main source of disturbance. We digitized the perimeters and entered the dates of all documented fires in the park since 1984. For plots that lay within these perimeters, the sample collection date was compared to the disturbance date; if the disturbance occurred after sampling, recent satellite imagery was used to assess the extent of disturbance. If conditions no longer resembled those present at the time of sampling, the plots were not used in modeling.[18]

2.2. Floristics

2.2.1. Debugging species lists

A fair number of inconsistencies in species nomenclature occurred in the ocular data, due to the use of at least two taxonomic references (Hitchcock and Cronquist 1973, Pojar et al. 2004). For the purpose of our analysis, we standardized the records by selecting the most frequently used name in

[17] Any disturbance impacting the park since this date will not be reflected in the map. Furthermore, areas disturbed within the time window spanned by the predictive imagery sources used (i.e., between summer 2015 and this date) may not be represented correctly.

[18] These plots were still used to develop floristic characteristics for the classification, despite no longer persisting in that condition.

the field datasets. We used Burke Herbarium (2020) to identify synonymies and to determine a standardized name for taxa where there was no prevalent name in the field datasets. Henceforth we refer to these standardized names, used in the field by NPS crews, as *field names*; we provide a crosswalk to Hitchcock and Cronquist (2018) in INR (2021b).

We systematically addressed problems involving confused taxa that resulted from consistent misidentifications during particular field collection efforts and from easily scrambled species names and codes. In some cases, resolving these and other thorny issues required that we refer to plot photos, field notes, location information and data from surrounding plots. Occasionally we fell back on the judgment of experienced botanists that a particular taxonomic record was unlikely. We considered the experience level of the field crew involved on such plots in making our decisions. Other cases were easier to resolve, such as recognizing that a *Eucephalus ledophyllus* record at OLYM was probably really *E. paucicapitatus*, based on their established range boundaries.

2.2.2. Expanding species lists

Plots collected under Protocols Y and Z emphasized collection of full-ocular data. Because these were the main protocols used for mapping plots at OLYM and NOCA, the relationship between floristics and map units (mapping associations or map classes) was very well characterized at those parks. In contrast, MORA had very few full-ocular plots, due to the use of the minimal Protocol X and the lower overall sampling effort beyond the classification phase. This created two related difficulties. First, it reduced the degree to which MORA plots were represented in the relationships developed between floristics and map units, which threatened to make the resulting map classification less applicable at MORA and thus less representative of the NCCN as a whole. Second, it made the MORA association calls—even at plots with full-ocular data—less reliable, because the distinct character of the vegetation there was not well-captured in the data.

In an attempt to address these concerns, we made additional efforts to improve the floristic completeness of many partial-ocular plots from MORA. For targeted plots in poorly sampled portions of the classification, a field botanist examined the plot photos and field notes, adding observed species and adjusting cover estimates. We tested the consequences of creating species lists entirely in this manner, comparing results to lists generated in the field at full-ocular plots. We found that on average three-quarters of the species with at least 1% cover were found. Automated placement of these office-created ocular records into an association[19] matched the full-ocular result at the association level at about 35% frequency and at the map class level at nearly 70% frequency. While field-collected full-ocular data are clearly preferable, these results represent a significant improvement over the level of field detail provided with most partial-ocular plots at MORA. We assigned all plots to association or map class manually, based not only on species cover estimates but on the cumulative weight of all available evidence.

In addition, we lowered the standards for labeling MORA plots as full-ocular, to better represent the park in map unit floristics. In general, if we determined that a plot's ocular data likely represented all

[19] Automated assignments were made using the species cover match tools discussed in **Section 2.2.4**.

prominent species and contained the most significant species in each vegetation layer,[20] we considered it full-ocular and used it in establishing the floristic characteristics of mapping associations and map classes. Despite these efforts, MORA remained poorly represented compared to the other parks, though the later incorporation of the forest legacy plots helped considerably (see **Table 2**).

2.2.3. Taxonomic treatment for floristic analyses

Plot-level species lists were used in all phases of this project. They were the primary source of data we used to assign plots to associations, to rework the mapping associations and form the map classes, and to describe the associations and map classes for users. However, the level of floristic detail captured varied by observer, collection effort, time of year and weather conditions, especially for uncommon, cryptic and ephemeral taxa. Additionally, the number of observations of many less common species fell short of the sample sizes needed to generate reliable statistics. To address these issues, we aggregated rarer species into groups to increase statistical strength and took other steps to reduce variability in floristic detail across plots. We defined a set of analysis taxa in which common and readily identifiable species were treated at species level, while less common or troublesome (cryptic or otherwise difficult) taxa were treated at genus level or as intermediate sub-genus groups defined by lumping species with similar habitats. Some infrequently observed taxa were dropped entirely from analyses involving plot-level comparison.

Troublesome species found in the field were often identified at the genus level.[21] Because our floristic analysis presupposed that the same taxonomic units were used across all plots, leaving these records at genus level would have required lumping the genus and sacrificing the species-level data collected across all other plots. To avoid this, we worked to link genus-level observations to a more specific taxon, particularly for common genera associated with diverse habitats. We accomplished this by creating sub-genus species groups with similar overall morphology, habitat requirements, distributional data and community affinities.[22] We then assigned genus-level field records to the sub-genus groups based on the weight of evidence at plots (e.g., elevation, topographic position, species co-occurrence matrices). Other less common species and genera were lumped to either the sub-genus or genus level, in order to gain necessary sample sizes for analysis.

Species of *Carex* and *Salix*, which are key indicators of several vegetation types, required the most attention. In these widespread yet difficult genera, the group formation process focused primarily on morphology and habitat requirements. As an example, unknown dwarfed alpine willows were coded as *Salix nivalis*+, which was defined to include *S. nivalis* as well as *S. petrophila* and *S. cascadensis*.

---

[20] We made this determination with reference to our own field experience and by comparing the species lists to available field photos and to other data collected nearby.

[21] A few records were identified at family or higher taxonomic levels; these were excluded from analysis.

[22] Treatment at this level required that we merge the species-level data collected at other plots into the same sub-genus categories, so determining appropriate categories was critical. Former NPS field botanists Matt Lee, Tynan Ramm-Granberg and Rachel Brunner were instrumental in this step.

This entailed losing the distinction between the three alpine species, but we deemed that far preferable to combining all of them with unlike lower elevation species such as *S. commutata* and *S. sitchensis.* Other genera, in which species were less clearly sortable by morphology and life zone, were treated primarily based on species affinity data, using the full plot database to develop co-occurrence relationships and sub-generic groupings.[23] Finally, some species that were often confused by field crews were also lumped (e.g., *Juncus parryi* and *J. drummondii* were lumped as *J. parryi*+).

Across 49 genera, 73 distinct sub-genus taxa were created in this manner, with genus-level occurrences assigned downwards to them and species-level occurrences lumped upwards into them. 65 other genera were treated at the genus level, lumping species-level occurrences up. These taxa, in addition to the species treated at species level, are cumulatively referred to as *SCM taxa* (see **Section 2.2.4**). **Table 2** in INR (2021b) identifies the SCM taxon used for each field-identified taxon.[24] Prior to publication, plant nomenclature was updated to match Hitchcock and Cronquist (2018); the resulting name changes are documented in **Table 1** of INR (2021b).

2.2.4. Floristic analysis tools

An enormous quality control effort was needed to bring consistency to the association calls across the more than 6,100 field plots that were available for use as model training data. We developed several floristic analysis tools to allow us to objectively evaluate and prioritize the review of association labels. The tools were also used to help guide the development of mapping associations (**Section 2.3**). They are briefly described below.

Species cover match tool

Associations, and the map classes we developed from them, are defined by their floristics, their physiognomic structure, and their position along multiple environmental gradients. We developed a tool called *species cover match* (SCM) to provide a quantitative representation of the degree of fit of a plot to the floristic and (to a lesser extent) structural aspects of a class, and to flag plots that were outliers within the class to which they were assigned. Generally, these resulted from field crews having encountered vegetation communities that had not been treated in Crawford et al. (2009), from mixed species lists due to heterogeneous plots combining multiple vegetation patches, from artifacts relating to hard breaks in the keys, or from differing crew interpretations of how significantly to weight different components of the association descriptions.[25]

---

[23] For example, we treated the genus *Arnica* as three taxa for analysis: a sub-genus group *A. latifolia*+, containing *A. latifolia*, *A. longifolia*, and a taxon identified in the field as *A. alpina*; another sub-genus group *A. mollis*+, containing *A. mollis*, *A. parryi* and a taxon identified as *A. amplexicaulis*; and a distinct species *A. cordifolia*. Genus-level records were assigned to one of the three based on species co-occurrence data.

[24] Henceforth in this report, both actual taxonomic species and the sub-genus groups defined here may be referred to simply as species, for simplicity.

[25] For example, the description for *Alnus rubra/Polystichum munitum* stated that "the herb layer is always dominated by *Polystichum munitum*," and also that it "occurs on upland slopes" and "is [a] result of succession after [disturbance]." On encountering a plot on an upland slope initiated by disturbance that lacked *Polystichum munitum*

SCM used the R **vegclust** package (De Cáceres et al. 2010) to compute the multivariate floristic distance of each of the 4,100+ full-ocular plots from the centroid of each class, as defined by the plot labels.[26] In order to more closely align the analysis with the emphasis on vertical stratification in the NVC, we weighted the cover values of each SCM taxon by a lifeform-specific[27] multiplier, applied to the transformed and standardized cover values.[28] To approximate the NVC's structural emphasis, we calculated total cover for each lifeform and for all vascular vegetation,[29] and incorporated those in the analysis as if they were additional taxa.

Partial species cover match tool

Class labels on partial-ocular plots could not be evaluated reliably using SCM and vegclust, because they didn't include true absence data (i.e., crews may have simply omitted a species). To evaluate these plots, we developed the *partial species cover match tool* (pSCM), which compared the SCM taxon cover estimates for partial-ocular plots to expectations derived from the class constancy and cover tables computed from full-ocular plots. The tool output a similarity metric between each partial-ocular plot and each class, and could be used in several different modes.

Three options were available to control the functioning of pSCM: *full mode* versus *partial mode*, *cover mode* versus *presence mode*, and *lifeform mode* versus *no-lifeform mode*. In full mode, pSCM penalized absences of taxa that were characteristically present in a vegetation class, while partial mode ignored these and so allowed for more missing information. In cover mode, cover estimates for a taxon that were significantly greater or less than the average cover for the class were penalized, while in presence mode only the presence or absence of a taxon was considered. In lifeform mode, lifeform totals were used in the similarity estimate, in addition to taxon cover estimates. Any

---

dominance but otherwise matched the description, some crews would emphasize the setting and decide it was a good enough match, while others would put more emphasis on the insufficient *Polystichum munitum* and choose another alternative. Since the key required 5% or more cover of *Polystichum munitum*, that might often have been used to resolve the question. Using the full species list to make these decisions results in many fewer such ambiguities.

[26] For this analysis, we transformed percent cover via a modified exponential equation (resulting in rapid changes of the transformed value in the indicative 2–10% cover range) to mimic breaks in the original association keys and allow the multivariate data-driven results to maintain as much compatibility with the keys as possible. We then standardized with respect to the mean and standard deviation of each species across all plots.

[27] We assigned all taxa to the following lifeform categories: broadleaf tree, conifer, tall shrub, standard shrub, dwarf shrub, forb, grass, sedge, rush, fern, fern ally, bryophyte and lichen (see INR 2021b).

[28] We used a multiplier of 2.0 for conifer and broadleaf tree species, and 1.5 for tall and standard shrubs. All other lifeforms had a multiplier of 1.0. In order to give more weight to taxa that were instrumental in defining the Crawford et al. (2009) associations, the multiplier for each SCM taxon was increased from its lifeform default proportionally to its maximum constancy in any mapping association, up to a maximum of 0.5 for taxa that were always present in an association.

[29] For purposes of lifeform totals, broadleaf trees and conifers were split into two vertical categories, GT5 (height over five meters) and LT5 (height less than five meters, or *regen*). Lifeform and total vascular cover were transformed using a sigmoidal curve to emphasize change in the region of 10% cover, in keeping with treatment of these thresholds in the original association keys.

combination of modes from the three options could be selected, allowing tailoring of the assessment to the amount of information available at a plot.

Differential indicators tool

Finally, we also developed a *differential indicators tool* (DIT) which we used to determine which of two classes was a better fit to a plot based only on the presence of the documented SCM taxa. For each present taxon, DIT calculated the ratio between its constancy in two selected classes. Each ratio was clamped at a maximum value of 10 before taking its square root. The transformed ratios were averaged across all present taxa and compared between the two classes, with the class giving the highest average ratio favored.

SCM, pSCM and DIT were all used to assist in determining the best calls at plots, depending on the sampling effort at the plot. SCM was primarily used during the earlier plot QC stages while we were still ironing out the mapping associations, while pSCM and DIT were used more in the later phases, especially at MORA where full-ocular plots were in short supply. The ability to label partial-ocular plots confidently was extremely helpful at increasing the available training data for modeling less common map classes at all parks.

2.3. Mapping associations and plot label QC

Although the following steps are written in sequential order, the processes occurred in tandem. The development of mapping associations and the quality control of plot association calls were strongly iterative processes. We have attempted to describe the steps with a minimal number of references to other parts of the process, but to some extent that has been unavoidable.

2.3.1. Mapping associations definition and floristics-based plot QC

Early drafts of the vegetation maps were based, with minor adjustments, on the vegetation alliances defined by NatureServe (2012), which in turn were based on associations defined by Crawford et al. (2009). Model error rates (see **Section 2.7.5**) and preliminary comparison of draft maps from MORA (in 2011) and OLYM (in 2013) to independent accuracy assessment data (see **Section 3**) indicated that the maps were falling well short of accuracy goals. As discussed in **Section 1.2.1**, it became evident through working with the training plots that making field calls based on dichotomous keys had resulted in a noisy dataset that may not always have correctly responded to the intentions of Crawford et al. (2009). Another source of error may have been the ongoing evolution of the classification itself during the fieldwork. Regardless, in some portions of the classification, the associations—as defined by the groups of plots assigned to them—lacked the needed floristic cohesion to support repeatable field identification and accurate mapping.

By early 2015, when mapping plot data collected at NOCA were delivered to INR, more than 4,100 field plots with reasonably complete species composition data were available across the three parks for use in floristic calibration. This significantly exceeded the information that had been available for the development of the earlier classification. We used the cumulative dataset to enhance the classification for floristic consistency and mapping purposes, creating a set of *mapping associations*. Despite their differences, the NCCN parks share many dominant plants and plant communities. We took advantage of this commonality, so each park benefitted from plots collected across the network.

We began by reviewing full-ocular plots with SCM. Plots that were significantly more similar to a different association than that to which they were assigned were examined to determine why their floristics differed from expectations. We checked field photos, plot descriptions, imagery and environmental setting, and changed the call to the association suggested by SCM if the balance of evidence supported that. For classes that were strongly defined by their vegetation structure (e.g., krummholz), we were more lenient in allowing floristic outliers to persist.

The process was applied iteratively: as plot QC continued, the analysis was occasionally updated, tightening the floristic groupings as the number of outliers was reduced in each cluster. In this manner, we refined the Crawford et al. (2009) associations while minimizing changes to their essential character. SCM was also used to suggest a best call at plots for which no confident call had been previously made.

We continued the revision process by eliminating problematic types from the mapping associations. Beginning with the original 311 upland and 50 wetland types, we removed (a) associations with fewer than two floristic calibration plots;[30] (b) associations distinguished from others based solely on total vegetative cover, either cumulative or in a single layer;[31] (c) associations named and defined based on the presence of a single common species (often a dwarf shrub such as *Vaccinium deliciosum* or *Juniperus communis*), regardless of the other vegetation present;[32] and (d) associations that were excessively heterogeneous in species composition (as represented in the floristic calibration plots), occurring in a variety of settings.[33]

We used SCM to reassign affected plots to the next most similar vegetated association, which was usually a very good fit. In addition to eliminating associations and merging their plots with similar types, we developed new mapping associations for groups of plots that were either poorly represented in Crawford et al. (2009) or had become badly tangled in the floristic calibration plots. These groups included dry shrublands, dry subalpine and alpine meadows, vegetation of talus slopes and avalanche chutes, riparian and wetland shrublands, and seral post-fire vegetation. We created the

---

[30] These had often been included in Crawford et al. (2009) based on literature from areas adjacent to the NCCN parks. We retained one association with only one plot, *Populus tremuloides/Cornus nuttallii*, because of its distinctiveness and the clear range limitation that prevents it from being more widespread in the parks.

[31] Two examples in Crawford et al. (2009) are the associations labeled as "bryophyte and lithomorphic sparse vegetation," keyed under a break based on the total vascular cover, and the three depauperate understory forest associations, keyed on overstory species and low understory vascular plant cover.

[32] Most plots assigned to these calls were small and represented a localized patch of the species in question. Generally these patches did not correspond to any meaningful landscape pattern, but simply reflected the stochastic dispersal and establishment processes of the single species, superimposed on a variety of background vegetation types.

[33] Typically, these associations—which were termed *catchalls* by field crews—resulted from key artifacts. They were recognized by their tendency to model with a variety of map classes, depending on the other vegetation present in addition to the species on which the key had focused.

new associations by clustering all plots assigned to an association in each of the groups with the R **vegclust** package (De Cáceres et al. 2010).[34]

Forests with depauperate understories provide a good example of the sorts of changes we made to the classification. These plots—usually in seral stands, but occasionally in older forests on valley bottoms—were originally lumped into associations based solely on the tree canopy species present, but we found these often did not model and map well together.[35] DIT and pSCM were helpful in making the best use of the understory floristic data, even if plants only occurred in trace amounts. For mapping purposes, the identities of the species present were much more important than how much ground they covered. For instance, we found that a trace amount of *Orthilia secunda* was a consistent indicator of the most common successional mid-slope silver fir association. Silver fir plots with equally sparse understories that lacked *O. secunda* typically had moist site indicators instead, and had closer floristic and modeling similarities to lush silver fir associations found on lower slopes. The plots simply represented unusually sparse manifestations of those usually lusher types.

Our classification efforts resulted in a total of 228 mapping associations in the large NCCN parks. Nielsen and Brunner (2021) provide descriptions, including floristic and distribution details, as well as more information about the process of creating the associations from the original classification.

2.3.2. Mapping associations refinement and model-based plot QC

We also prioritized examination of individual plots using model results to identify plots that modeled better as an association different than their current assignment.[36] Plots that modeled poorly had often been noted as problematic by the field crew and were generally in heterogeneous areas, in very small patches, or had mismatched structure and floristics (frequently due to disturbance; e.g., a forest that had experienced a blowdown event and was now dominated by shrubs, but with understory species more typical of a forest). Other plots that modeled poorly had been mislocated due either to extreme GPS error or data entry errors; there was considerable feedback between association-level modeling and the spatial QC described in **Section 2.1.2**.

An occasional outcome of plot-level model-based QC was a decision that a plot should not be used in modeling because of a poor match to any association, an uncertain location, or both. These plots were

---

[34] We log-transformed raw percent cover data for each SCM taxon and normalized across sites using the decostand function in the R **vegan** package (Oksanen et al. 2019) before using k-means clustering in **vegclust**. We experimented with the number of output clusters until the results captured a similar level of detail to that used elsewhere in the classification.

[35] Depauperate conditions occur in the stem-exclusion phase of a range of successional forest types, and can persist for over a century in the Pacific Northwest (Agee 1993, p.193). Thus, seral forests can be impossible to place definitively into classifications relying on understory species composition, and can be easily confused with very distinct valley-bottom stands that are similarly depauperate (Franklin et al. 1988, p.126).

[36] We did this by creating random forests models (see **Section 2.7.5**) at the association level, examining the cumulative out-of-bag error associated with each plot, and noting the alternate associations with which it was most frequently confused.

often still useful in refining the classification's approach to disturbance or in identifying range extensions of associations known primarily from another park. Throughout the process, we incorporated these observations into refined descriptions of the structure, setting and range of each mapping association.

2.3.3. Final plot check with a hybrid assemblage labeling tool

In the above QC steps, we considered floristic and modeling similarities separately and only examined plots that failed to pass some test by a significant threshold. After development of the mapping associations had been completed, we used a final check—the *hybrid assemblage labeling tool* (HALT)—which considered the floristics and modeling analyses simultaneously to spot instances where both pointed in the same direction, but perhaps at a lower level of certainty. HALT enabled us to detect and reassign about 50 plots to an association that was a better overall fit.

The QC process, while lengthy, accomplished several critical steps toward development of the map classification and the associated map: (a) development of an association-level classification with high internal cohesion in both floristics and modeling tendencies; (b) development of clear descriptions of floristics, structure and setting for those associations; and (c) allowing maximum use of all plot data by improving the consistency of association calls on all plots, and particularly by assigning reliable calls to partial-ocular plots.

2.4. Map classification

2.4.1. Development of vegetated map classes

Building crosswalk

The low accuracies of early draft maps indicated that changes to the alliance concepts were needed. A crosswalk to combine mapping associations into floristically cohesive and mappable entities provided the structure around which revisions were organized. We used the draft alliances from NatureServe (2012) and their relationship to the associations in Crawford et al. (2009) as a reference point during the revision process.

Our goal in this process—described in greater detail in Brunner et al. (2017)—was to minimize class confusion, both during field interpretation and in the map. Our approach was data-driven, using a quantitative proxy for each of these confusion types. As a proxy for field confusion, we used floristic similarity, since the more floristically similar two classes are, the less likely field observers will agree on the correct label for a plot. SCM, described in **Section 2.2.4**, provided an easy way of quantifying this at the plot level. To represent map confusion, random forests model confusion was clearly the appropriate proxy, as that was the means by which we planned to produce the map.[37] The main constraint we placed on the process was to follow the NVCS protocol of a many-to-one crosswalk between mapping associations and map classes, in which each association was a member of a single

[37] We quantified model confusion as the out-of-bag error rate for a plot in a model attempting to discriminate between a pair of associations, built from the plots assigned to either of them. The R **randomForest** package (Liaw and Wiener 2002) provides this information as an optional 'votes' table output. See **Sections 2.5–7** for background in the modeling process.

map class. In order to foster consistent map class definitions across NCCN parks, we aimed to use the same crosswalk for each of the mapping projects.

We began by identifying common and distinct mapping associations, emphasizing those that represented the cores of alliance concepts from NatureServe (2012). We used these as seeds for initializing map classes. If possible, we selected associations that were present at all NCCN parks in order to provide a common thread. If this was not possible, and we were confident about the relatedness of floristically dissimilar associations, we occasionally initialized map classes using a different association at each park. We did this in the case of vegetated balds, which are characterized by a common structure and setting but whose constituent species vary significantly with geography.

We then used an agglomerative process to grow the map classes from their seeds. At each step, we computed the level of floristic and modeling similarity (termed *joint similarity* hereafter) between each unassigned association and each nascent map class, by aggregating plot-level data.[38] We found the association–map class pair with the greatest pairwise joint similarity and joined them by assigning the association to that map class in the crosswalk. Association–map class similarities were recalculated after each assignment, and the next most similar pair found. The process resulted in maximizing within-class similarity and minimizing between-class similarity, allowing more confident discrimination in the field and more reliable mapping.

Early in the process, assignments were easy because many associations clearly belonged together based on both floristics and modeling. The decisions became more difficult later. When we encountered associations whose floristic and modeling tendencies pointed to different map classes, we emphasized the floristics, unless some overriding structural or setting-based criterion was available to assist in field identification. When different patterns of similarity were observed at different parks, we made our decision based on the park where the majority of association plots occurred. Occasionally we went back to the plot data to unravel problems.

Associations that fit poorly to existing map classes were added as new classes if they represented a distinguishable concept and had enough plots to support modeling. It then occasionally become apparent that other associations that had already been assigned had a stronger affinity for the new class. The iterative process continued until all associations had been assigned.

Refining crosswalk

After the crosswalk was formed, we recalculated association-wide model similarity to each full map class, again from plot-level data. We re-examined associations that were a better fit to a map class other than the class with which they had been lumped. We often found that this mismatch arose from plots that were floristically distinct from most others assigned to the association. These outliers usually were easily recoded on an individual basis to an alternate association, but in several cases we found associations that contained a full subset of plots that were similar to each other but distinct

[38] We computed similarity by aggregating plot-level data at each step because some plot-level QC was ongoing during this process and this prevented our needing to run random forests again with every change.

from the rest. We formed new associations with these plot subsets and moved them to a different map class. Nielsen and Brunner (2021) includes several examples of these new associations.

In many cases, map class occurrences were confined to only one[39] or two[40] of the NCCN parks, which presented no challenge to the crosswalk since the constituent associations were also absent. However, occasionally a map class was present in a park, but with too few plots from which to construct a model of its distribution. In these cases, we lumped the constituent associations with the most similar map class that was mappable at that park.[41] These are the only cases where the crosswalk between association and map class differs between parks.

In general, the outcomes of the crosswalking process confirmed our belief that unless aberrant vegetation structure was present, modeling tendencies and full floristic character tended to track each other extremely well. By maximizing the floristic distinctions between the map classes, we simultaneously created a highly mappable classification. The description for each map class in Nielsen et al. (2021c) contains a list of its component associations.

2.4.2. Development of other map classes

Natural sparse and abiotic map classes

Classification and mapping efforts were primarily focused on vegetated communities, but sparsely vegetated and abiotic areas occupy a large proportion of each NCCN park. To fill these areas of the map, we developed map classes that were simple for field crews to discriminate but would provide useful habitat context. We developed a "rock-dominated" set of map classes distinguished by the geomorphological origin of the mineral substrate, including R70–COASTAL BARREN, R71–ALLUVIAL BARREN AND DEBRIS-COVERED ICE, R72–COLLUVIAL BARREN and R73–BEDROCK BARREN. We also developed an "H_2O-dominated" set of classes composed of W80–SALT WATER, W81–FRESH WATER and W82–EXPOSED SNOW AND ICE.

Disturbed and cultural map classes

Several other classes were created to handle areas of uncertain vegetation impacted by recent fires or anthropogenic disturbance (the latter most often in the mapped buffer around the park). In this category were M92–BURNED WITH UNCERTAIN VEGETATION, M93–TIMBERLAND WITH UNCERTAIN VEGETATION, M94–DEVELOPMENT and M95–ROADS IN PARK. Details on discriminating these and the preceding types are contained in the map class descriptions (Nielsen et al. 2021c).

---

[39] Examples included H63–ALPINE BUCKWHEAT PUMICE VEGETATION (MORA only), C02–REDCEDAR, LABRADOR-TEA, SLOUGH SEDGE AND SPHAGNUM BOG (OLYM only), and C22–SUBALPINE LARCH WOODLAND (NOCA only).

[40] Examples included C03–SITKA SPRUCE, WESTERN HEMLOCK AND WOOD-SORREL FOREST (absent at NOCA), H57–GREEN FESCUE DRY MEADOW (absent at OLYM), and H52–COW PARSNIP MEADOW (absent at MORA, at least at mappable patch size).

[41] Examples included B33–UPLAND RED ALDER, BIGLEAF MAPLE AND CONIFER FOREST (treated as S45–VINE MAPLE SHRUBLAND at MORA) and H56–SUBALPINE SUMMER-DRY GRASS-FORB MEADOW (treated as H57–GREEN FESCUE DRY MEADOW at NOCA).

2.4.3. Resulting map classification

The map classes present in the OLYM map are summarized in **Table 3**. See Nielsen et al. (2021c) for detailed map class descriptions and an explanation of the map class name coding system.

Table 3. Map classes present in the OLYM map the other NCCN park maps in which they appear, and the number of training plots called to each at OLYM.

Map class code and full name	Other parks	Plot count
C00–Sitka spruce, western hemlock, salal and sword fern forest	–	91
C01–Redcedar, western hemlock, salal and evergreen huckleberry forest	–	133
C02–Redcedar, Labrador-tea, slough sedge and sphagnum bog	–	27
C03–Sitka spruce, western hemlock and wood-sorrel forest	MORA	84
C04–Moist western hemlock, Douglas-fir and foamflower forest	MORA, NOCA	51
C05–Western hemlock, Douglas-fir and sword fern forest	MORA, NOCA	106
C06–Western hemlock, Douglas-fir and salal forest	MORA, NOCA	145
C08–Olympic dry Douglas-fir forest	–	55
C10–Moist silver fir, western hemlock and foamflower forest	MORA, NOCA	100
C11–Mesic silver fir and western hemlock forest	MORA, NOCA	105
C12–Silver fir, hemlock and Alaska blueberry forest	MORA, NOCA	44
C13–Mountain hemlock, silver fir and Cascade azalea forest	MORA, NOCA	51
C14–Silver fir, big huckleberry and beargrass forest	MORA, NOCA	12
C15–Lodgepole pine and Douglas-fir woodland	MORA, NOCA	18
C20–Subalpine fir and Sitka valerian forest and woodland	MORA, NOCA	93
C21–Mountain hemlock, subalpine fir and heather woodland	MORA, NOCA	106
C24–Olympic subalpine fir and lodgepole pine woodland	–	26
C26–Conifer krummholz and treed cliff	MORA, NOCA	37
B30–Successional gravel bar shrubland	MORA, NOCA	18
B31–Broadleaf riparian and swamp forest	MORA, NOCA	108
B32–Coastal red alder forest	–	35
B33–Upland red alder, bigleaf maple and conifer forest	NOCA	38
S40W–Low elevation shrub-dominated wetland	MORA, NOCA	10
S41W–Subalpine willow wetland	MORA, NOCA	36
S42–Sitka willow riparian shrubland	NOCA	9
S43–Sitka alder shrubland	MORA, NOCA	56
S44–Thimbleberry shrubland, tall forbs and bracken fern	NOCA	16
S45–Vine maple shrubland	MORA, NOCA	12
S47–Successional huckleberry shrubland	MORA, NOCA	26
S48–Subalpine heather shrubland	MORA, NOCA	61
S49–Alpine heather shrubland	MORA, NOCA	15
H50W–Lowland marsh and meadow	MORA, NOCA	27

Table 3 (continued). Map classes present in the OLYM map, the other NCCN park maps in which they appear, and the number of training plots called to each at OLYM.

Map class code and full name	Other parks	Plot count
H51W–Subalpine herbaceous wetland	MORA, NOCA	28
H52–Cow parsnip meadow	NOCA	37
H53–Showy sedge and Sitka valerian meadow	MORA, NOCA	81
H54–Moist talus vegetation	NOCA	9
H55–Olympic aster scree meadow	–	28
H56–Subalpine summer-dry grass-forb meadow	MORA	25
H58–Bedrock balds and sparsely vegetated forest openings	MORA, NOCA	17
H60W–Black alpine sedge wetland	MORA, NOCA	17
H61–Spreading phlox and common juniper dry meadow	–	68
H62–Alpine sparse herbaceous vegetation	MORA, NOCA	18
R70–Coastal barren	–	27
R71–Alluvial barren and debris-covered ice	MORA, NOCA	57
R72–Colluvial barren	MORA, NOCA	74
R73–Bedrock barren	MORA, NOCA	48
W80–Salt water	–	24
W81–Fresh water	MORA, NOCA	72
W82–Exposed snow and ice	MORA, NOCA	21
M90–Ruderal meadow	–	12
M92–Burned with uncertain vegetation	MORA, NOCA	–
M93–Timberland with uncertain vegetation	MORA, NOCA	–
M94–Development	MORA, NOCA	–
M95–Roads in park	MORA, NOCA	–

Development of the NVC for the Pacific Northwest continued on a somewhat parallel track to ours, as we worked on finalizing the mapping associations and map classes presented here. We compared the relationship of our mapping associations and map classes with the hierarchical placement of the related associations in the most recent NVC update, USNVC (2019). At the NVCS group level, there is good correspondence, with our map classes mostly composed of associations that are members of a single group, or of an amalgam of associations from groups that are poorly represented in the project area and do not overlap with other map classes. There is less congruence at the alliance level, with one cause being that our map classes are generally less beholden to dominance and encompass a broader range of indicator species. Forests are somewhat more finely delineated in map classes than the current NVC alliances, but non-forests are a bit more coarsely lumped. Structural characteristics appear to be more important in distinguishing forested map classes than the corresponding alliances, but less important in distinguishing dwarf shrubland and herbaceous map classes.

2.4.4. Descriptions

Summary and setting narratives

The map class summary and setting paragraphs in Nielsen et al. (2021c) were compiled from plot-level floristics, vegetation structure data, summarized environmental variables and expert knowledge. We edited the narratives to reflect park-specific characteristics and added observations based on the final map and plot data. Representative plot photos were selected for each class; these were generally obtained from the park in which it was most common.

Floristics tables

We generated constancy and cover information for each of the resulting map classes, based on the complete set of full-ocular plots across the three parks. Because of the large number of records that were uncertain at the species level, we used the SCM taxa described above (and documented in INR 2021b) rather than species as the taxonomic units for the analysis. We used the tables generated to assign descriptive names to the map classes. Each map class description in Nielsen et al. (2021c) contains a condensed version of the constancy and cover results.

Indicator species analysis

Although we relied on the tools described above (SCM, pSCM and DIT) to reliably discriminate between map classes on the basis of plot floristics, those tools will not be available to field users unless they are carrying a mobile device. As an alternative, we created a park-specific list of indicator species that are helpful for distinguishing each pair of map classes, which we have included in Nielsen et al. (2021c) for pairs that are likely to be occasionally confused.

We derived indicators from the constancy and cover data. *Presence indicators* are SCM taxa that are significantly more likely to be present in one of the map classes than in the other, based on the constancy tables. We rated the strength of presence indicators by the constancy ratio between the two classes and put those ratings on a comparable scale for both sides of each map class pair. *Cover indicators* are taxa that are likely to occur in significantly greater abundance in one of the classes than in the other, based on the cover tables. We prioritized listing taxa that occur reasonably often in the favored map class, but in some cases only less common taxa are good indicators. For this reason, we listed a significant number of indicators. Lack of presence of an indicator is not evidence against a map class; however, absence of taxa listed as occurring at high frequency in the floristics table for a map class can be construed that way.

2.4.5. Key

We have mentioned several times the difficulties we encountered with field plots that had been assigned to association based on a dichotomous key (see **Section 1.2.1** and elsewhere). However, although "the key is not the classification" (Crawford et al. 2009), it is where a typical user will start. In our map class key (Brunner et al. 2021), we aimed to provide as much help to a field user as possible without leading them astray by oversimplification. We strongly urge users who have keyed to a map class to carefully consult the map class description, including the indicators for closely-related alternate classes.

Much of the key was built using automated methods such as multivariate hierarchical clustering via **hclust** (R Core Team 2018) and classification trees via **rpart** (Therneau et al. 2015). Because setting and structural characteristics are easiest for a non-botanist to identify, we prioritized them in the key where possible, mostly at higher levels. At the lower levels (e.g., within conifer forests), the breaks were mostly determined by floristics. We transformed plot species composition information into binary true/false characteristics based on presence, prominence, dominance of individual species and functional groups (e.g., broadleaf trees, total vascular cover), relative abundance (e.g., cover of *Acer circinatum* significantly greater than that of *Alnus viridis*), and quantifiable setting variables (e.g., south-facing). We used **rpart** to determine the optimal structure and best key break variables based on pools of samples drawn from the full-ocular plots, assuming these decisions would translate to new plots encountered in the field.

To help keep users from taking a wrong turn based on a single criterion, we added additional floristic and setting-based characteristics at most breaks to lend additional confidence. The additional criteria were pulled from surrogate variables in the classification tree, from a break-specific indicator species analysis using the R **indicspecies** package (De Cáceres and Legendre 2009), and from setting and structure notes. After each break was written, it was applied to the current plot pool and the resulting subdivided pool was fed into the next break. We minimized misclassification by only including criteria that correctly classified 95% or more of the plot pool entering the break; we tried to find a way to shepherd the misclassified plots home later in the key. The key was validated in the office with over 200 field plots per park and was also briefly tested in the field.

2.5. Independent data selection and pre-processing

We used an implementation of the random forests machine learning algorithm (Breiman 2001) to predict map class from field training data (discussed in **Sections 2.1–4**) and wall-to-wall independent predictor data. We used several broad categories of predictor data: (a) four separate years of aerial imagery from the National Agricultural Imagery Program (NAIP); (b) satellite imagery from the Landsat-8 and Sentinel-2 satellites, collected during multiple distinct seasons; (c) topographic and hydrologic metrics developed from lidar data and standard digital elevation models; (d) climate normals over the period 1981–2010; and (e) vegetation canopy information from lidar data. Several other types of potential predictor data—soils, surface geology, and geological landform information, infrastructure development locations, and maps of fire history—were considered for use and ultimately rejected. These layers had poor spatial registration or were incomplete or inconsistent over the project area. We felt their use would likely result in mapping artifacts and would add little predictive power, since correlated information was available already in the other predictors. **Table 4** summarizes the data sources with their spatial resolutions and the dates to which they apply. The selection, acquisition and pre-processing of the data are described in the following sections.

Table 4. Sources of predictive modeling layers Multiple satellite imagery collections were necessary to cover both the main interior section of the park and the coastal strip.

Data type	Spatial res (m)	Data source	Applicable timeframe(s)
4-band color-infrared aerial imagery	1	National Agriculture Imagery Program, State of Washington	2009, 2011, 2013, 2015
Historic mid-summer image (main), Landsat-5	30	USGS (2019a)	Aug 1, 1986
Historic mid-summer image (coastal), Landsat-5	30	USGS (2019a)	Jul 17, 1984
Current mid-summer image (main), Sentinel-2	10, 20	USGS (2019a)	Aug 4, 2017
Current mid-summer image (coastal), Landsat-8	30	USGS (2019a)	Jul 23, 2015
Current minimum-snow image (main), Landsat-8	30	USGS (2019a)	Aug 7, 2015
Current minimum-snow image (coastal), Landsat-8	30	USGS (2019a)	Sep 6, 2014
Elevation	10	USGS (2019b)	–
Climate normals	~800	PRISM Climate Group (2019)	1981–2010
Lidar bare earth and highest hit elevation	3	Houze et al. (2017)	Sep 2014

2.5.1. Aerial imagery

We acquired 4-band color infrared NAIP imagery as uncompressed quarter quads from four separate collections, in 2009, 2011, 2013, and 2015. The 2015 imagery was the main recent data source allowing mapping at 3-meter resolution. However, deep shadows which lowered mapping accuracy often occurred north of steep slopes (**Figure 6** illustrates this with an excerpt from NOCA). Because the shadow locations varied between image collections, we mitigated this problem by combining the best-illuminated portions of each into a mosaic.

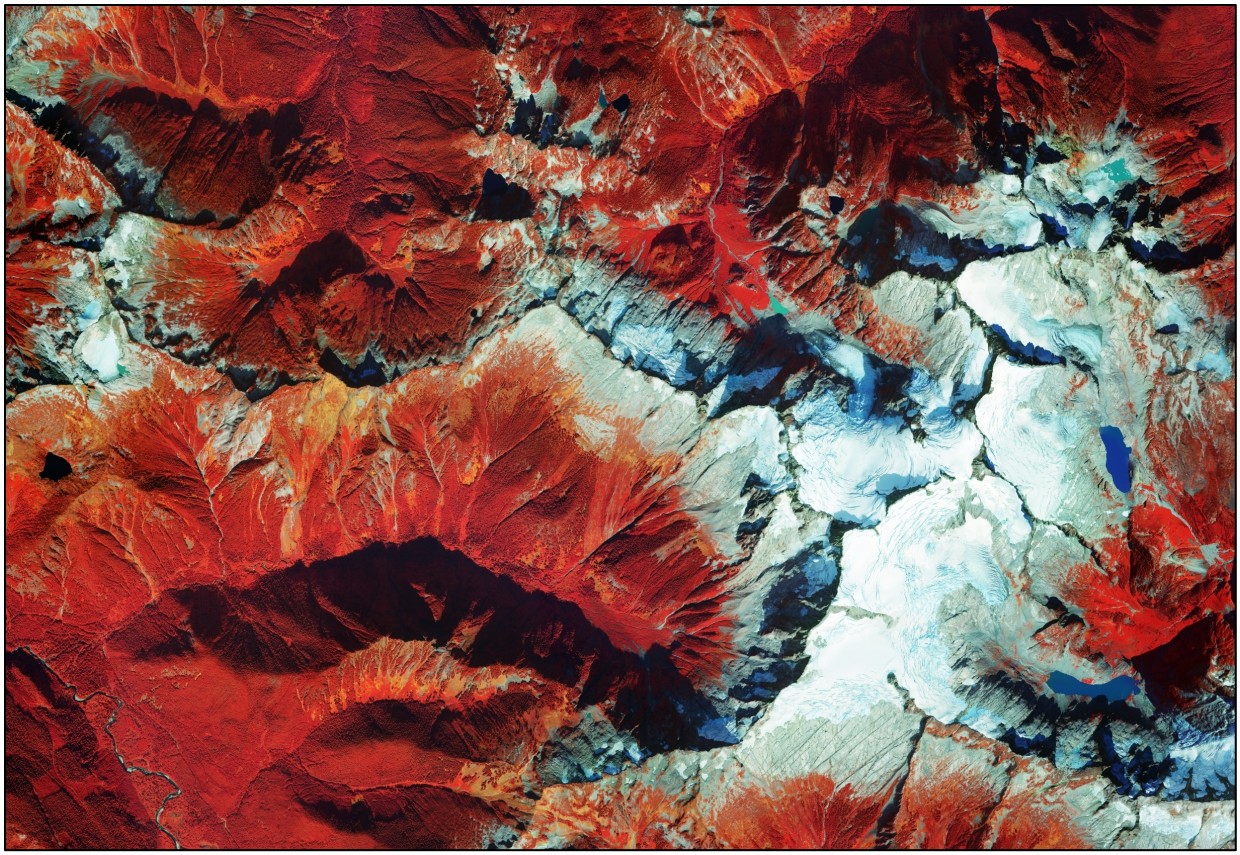


Figure 6. Aerial imagery with untreated shadows The deep shadows seen in this 2015 NAIP imagery would interfere with accurate mapping unless treated.

Correction of 2015 imagery

We mosaicked the uncompressed quarter quads from each NAIP collection and generated aerial imagery metrics (**Section 2.6.1**). Making use of topographic information (**Section 2.6.3**), we then built a predictive model to identify shadows in the 2015 imagery by digitizing shadow and non-shadow training data, identifying shadows using a random forests model, and iteratively selecting additional training data to home in on problem areas. When satisfied with the results, we converted the shadow mask to a shapefile and buffered each feature by a variable distance, using a formula that yielded a buffer area roughly proportional to the size of the feature. Our hypothesis was that over a given region, the histogram of pixel values for each image band within corrected shadows should resemble that within the adjacent unshaded areas. We broke the project area into overlapping tiles, derived a crosswalk between shadow pixel values and corrected values based on matching the shadow and buffer histograms, and applied this to all shadow pixels. **Figure 7** illustrates a resulting corrected image.

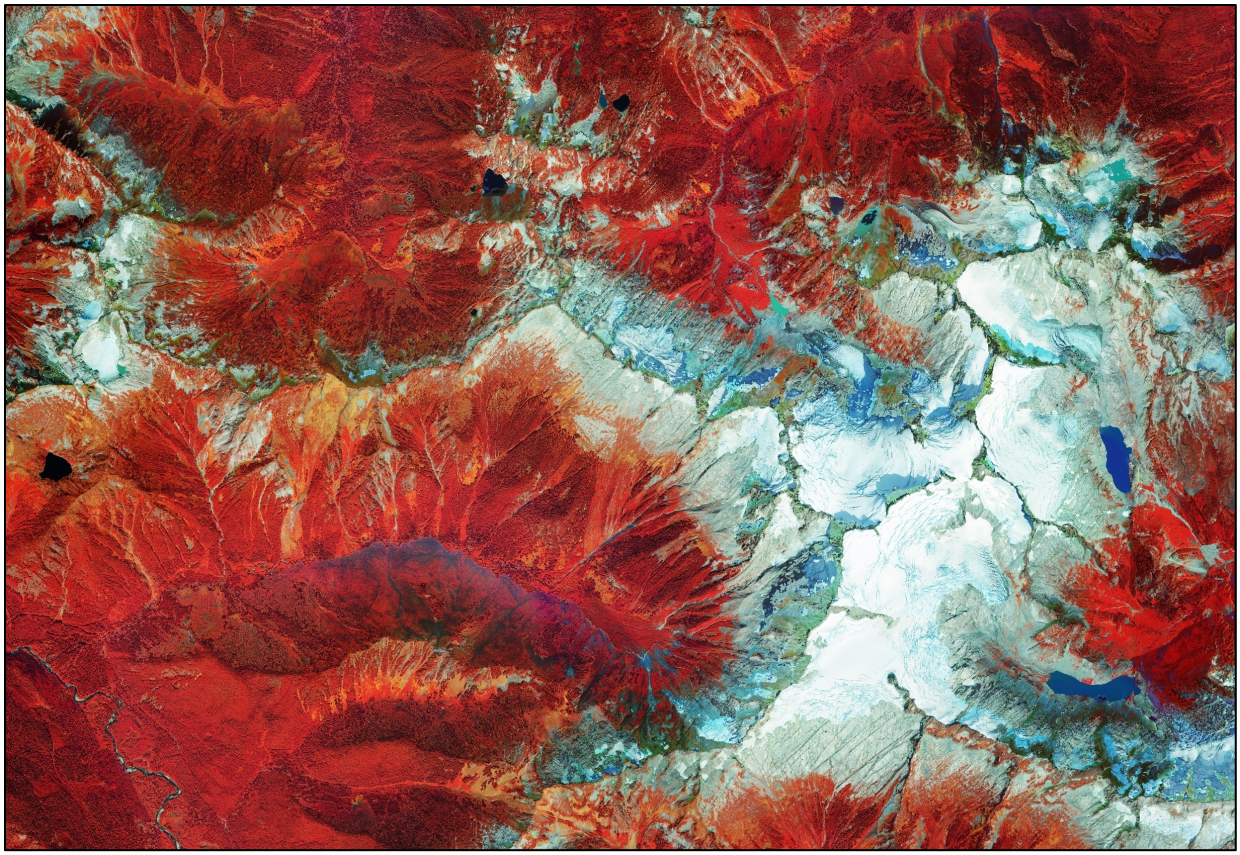


Figure 7. Aerial imagery with histogram-matched shadows Here the shadowed areas have been matched to the surroundings.

Image merging and correction

Although shadow pixel values in the corrected 2015 image showed reasonable correspondence to the underlying land cover, lack of direct illumination resulted in a major reduction in local variance which could not be corrected. Because of the importance of high-resolution texture in accurate identification of land cover types (see **Section 2.6.1**), we incorporated an additional method of shadow treatment. We applied the model generated from 2015 imagery to the other imagery years, yielding shadow/non-shadow masks for each year. These masks were used to produce a merged image by selecting the first non-shadowed year from the sequence (2015, 2013, 2009, 2011) subject to the condition that if a given pixel was located within the digitized fire perimeters (**Section 2.1.2**), only imagery collected after the fire year could be selected. The year 2011 had lowest priority in the merge sequence because high snowpack that year obscured the ground and delayed vegetation development at high elevations through much of the summer. The merged image replaced many of the shadows in the 2015 image with illuminated data from other years. Although the spectral characteristics differed somewhat from year to year due to the lack of radiometric normalization in NAIP data, we felt that for modeling purposes the result was far superior to leaving the shadows untreated. The areas that were shadowed in all imagery years were corrected using the procedure applied to the 2015 imagery above. The resulting image, suitable for generating texture metrics (**Section 2.6.1**), is shown in **Figure 8**.

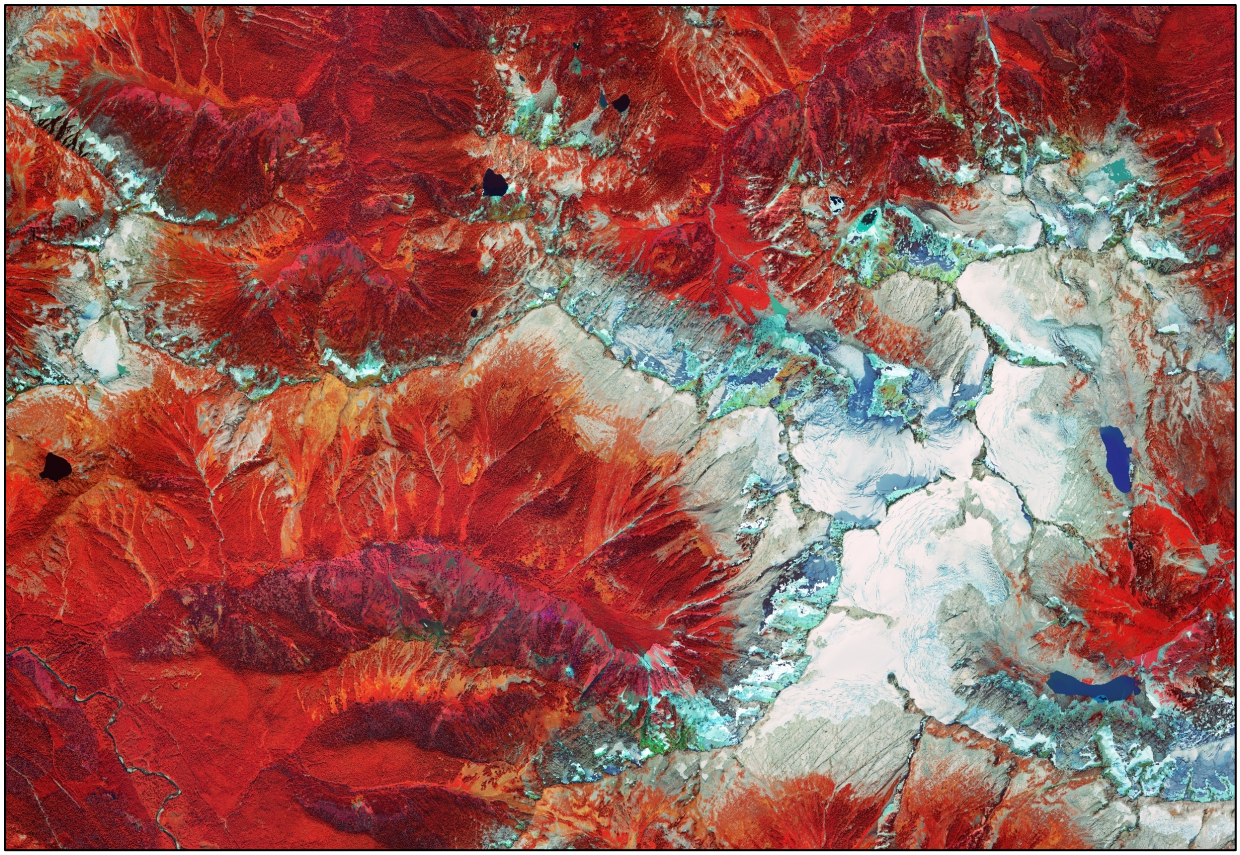


Figure 8. Aerial imagery merged across years Here the shadowed areas have been filled with data from the 2009–13 images. Areas shadowed in all images were matched to the surroundings.

2.5.2. Satellite imagery

We searched the image archive at GLOVIS (USGS 2019a) for cloud-free Landsat-5, Landsat-8, and Sentinel-2 images collected between early June and late September in all years since 1982. Images from outside that seasonal window were mostly snow-covered or had very low sun angles and were not useful for vegetation mapping. In the map training fieldwork phase, a Landsat-5 image collected on July 15, 2009 was used to guide sampling and produce field map sheets. A midsummer Sentinel-2 image collected on August 4, 2017 (**Figure 9**) was used as the primary satellite image for modeling. This was supplemented on the coastal strip by a Landsat-8 image collected on July 23, 2015. A Landsat-8 image collected on August 7, 2015 was also used in modeling, since it had the lowest snow cover of any available images, allowing more effective mapping of higher elevation areas. This was supplemented on the coastal strip by a Landsat-8 image collected on September 6, 2014. Finally, a Landsat-5 image collected on August 1, 1986 served as the starting point for historic change detection over the intervening time period. This was supplemented on the coastal strip by a Landsat-5 image collected on July 17, 1984.

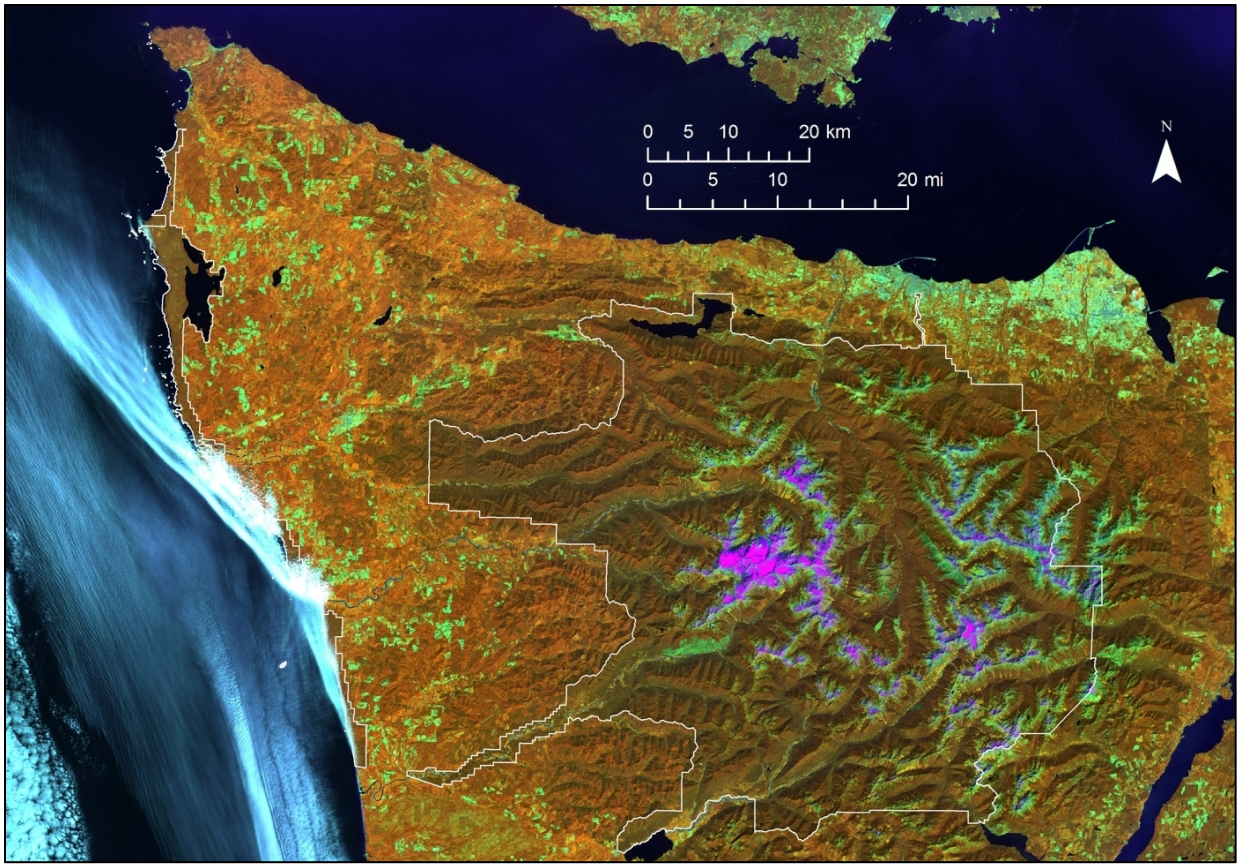


Figure 9. Base Sentinel-2 satellite image collected August 4, 2017. The intensity of shades of red represents near-infrared reflectance from the land surface, green intensity represents mid-infrared reflectance, and blue intensity represents red reflectance. The marine layer over the southern half of the coastal strip necessitated merging with a different image to provide complete coverage of the park.

All satellite images were converted to at-sensor reflectance (e.g., Chander et al. 2009), and a simple dark object atmospheric correction (Chavez 1988) was applied to approximate surface reflectance. We developed a novel process for spatial coregistration of the satellite images with the elevation dataset. We began by coregistering the minimum-snow image—which showed the greatest illumination contrast due to its acquisition at a time of relatively low sun elevation angle—to the elevation data. A cosine(i) image of illumination intensity at the time of image acquisition was created based on local slope and aspect; it served as a reference for aligning the satellite near-infrared band using the ERDAS Imagine Autosync tool. We then coregistered the midsummer image to the minimum-snow image using their respective near-infrared bands. The resulting coregistered images were resampled via cubic convolution to a common extent and pixel size. The coastal and inland images for each time frame were radiometrically normalized using reduced major axis regression (e.g., Cohen et al. 2003) on their overlap areas and then were joined.

The satellite images were then topographically normalized to reduce the effect of variable illumination on at-sensor reflectance. We did this via a modified version of the stratified c-correction method (Twele et al. 2006), using the normalized difference moisture index (NDMI; Wilson and

Sader 2002) for stratification of pixels into distinct correction groups. The normalization process reduced the effects of shading, causing individual land cover types to exhibit more consistent reflectance across the image, regardless of slope and aspect (**Figure 10**).

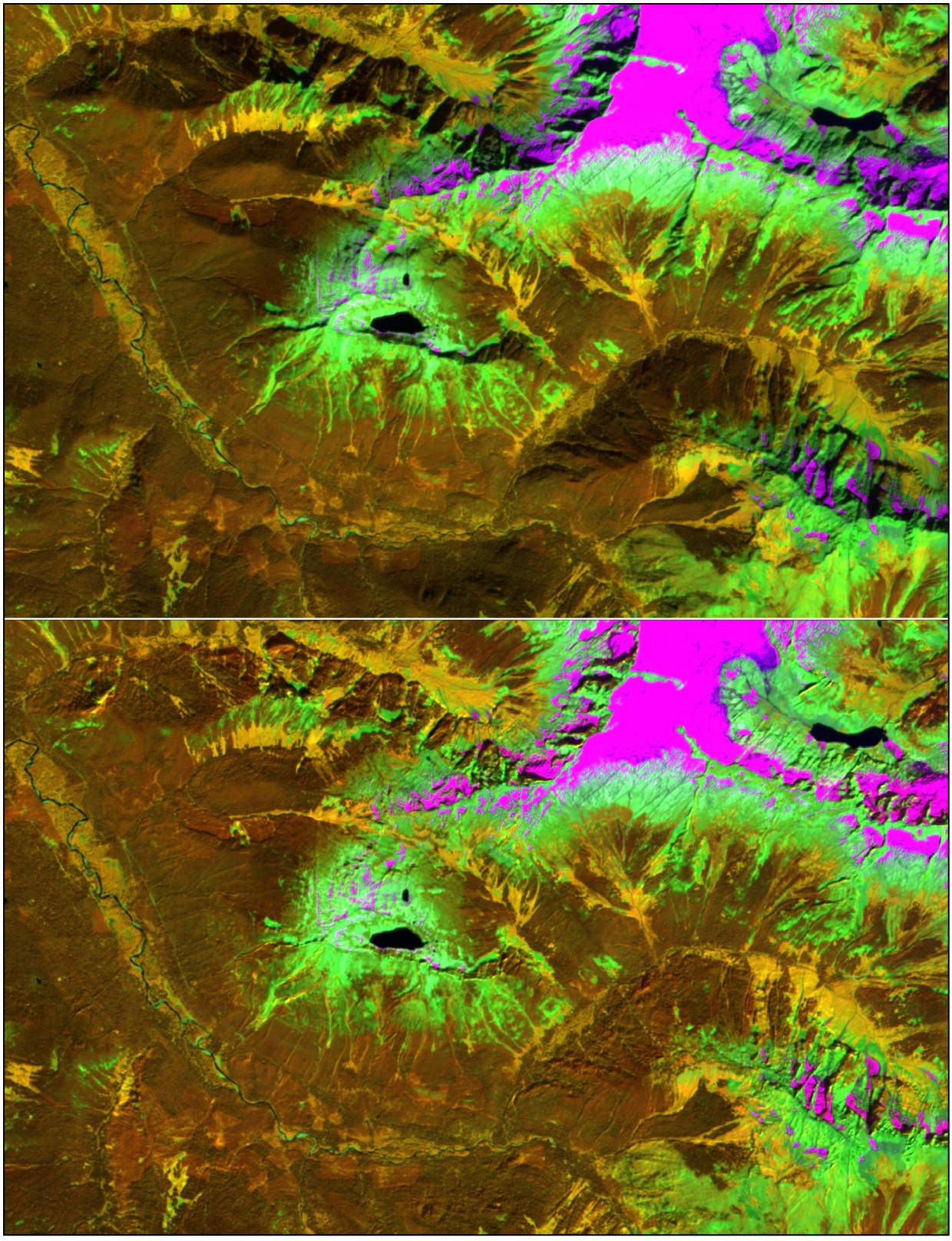


Figure 10. Sentinel-2 image before and after topographic normalization(upper and lower images respectively). Snow and ice appear pink, sparsely vegetated areas green, broadleaved trees and shrubs yellow, and conifers reddish-brown.

2.5.3. Elevation and climate data

We downloaded 10-meter resolution elevation data for the project area from the 3D Elevation Program (3DEP; USGS 2019b). We also downloaded a range of 30-year monthly climate normals at approximately 800-meter resolution from the PRISM Climate Group (2019), including January, April, July and October precipitation, minimum and maximum temperature, mean dew point temperature, and maximum vapor pressure deficit. For processing efficiency, the elevation data were converted to integer format using a vertical unit of 0.25 feet. The climate data were clipped to the project area, reprojected and resampled to 30-meter resolution using bilinear interpolation.

2.5.4. Lidar data

Lidar data were collected in September 2014 and preliminarily processed by the Olympic Mountains Experiment (OLYMPEX; Houze et al. 2017).[42] The collection covered nearly the entire park; only the coast and a small area near the southeastern park boundary were omitted. We obtained two raster datasets: the first return or *highest hit elevation*, and a processed "lowest hit" (Painter et al. 2016) referred to here as *bare earth elevation*. We converted both files from floating point to integer format with a vertical unit of 0.25 feet and aligned them with the other high-resolution datasets.

The main target of this lidar collection was above snowline; relatively sparse point clouds, or perhaps an attempt to compensate for them, resulted in artifacts in the bare earth layer in dense forests. We identified these artifacts using an elevation and slope-based anomaly filter. Anomalous areas were filled based on nearby areas using **Elevation Void Fill** in ArcGIS 10.3 and a 3x3-cell median filter was used to reduce the small artifacts that remained.

Vegetation height was calculated by subtracting the smoothed floating point bare earth elevation from the highest hit and converting to integer format with a vertical unit of 0.01 feet. An artifact in the vegetation height layer was observed in steep areas, where height values proportional to ground slope were found, even in areas completely barren of vegetation. Some steep cliffs resulted in height artifacts of 100 meters or more. Even on slopes of 30 degrees or less, invalid height data on the order of several meters occurred often. These artifacts presented a serious challenge due to the importance of vegetation height in high-resolution mapping. We identified erroneous height data by using a random forests model, trained from 200 anomalous pixels, to identify similar pixels across the park. We used neighborhood slope, elevation, and normalized difference vegetation index (NDVI; Rouse et al. 1974, Tucker and Sellers 1986) calculated from 2015 NAIP as predictors. The model achieved an accuracy of 93%, with the primary commission errors falling in small tree islands in the subalpine parkland. Vegetation height values of the detected artifacts were recoded as missing data.

2.6. Predictive metrics

We used spatial contextual information, variable transformations, and noise minimization techniques to produce predictive metrics with stronger relationships to vegetation patterns than the raw independent data. The metrics fall into six main categories: metrics derived from aerial imagery,

[42] OLYMPEX was a ground validation field campaign by the University of Washington and the National Aeronautics and Space Administration.

satellite imagery, topographic information, hydrologic information, climate data, and vegetation canopy characteristics. Each metric category is followed by a table detailing the predictive metrics produced from that data source. For each metric, we give an *effective resolution*. This combines characteristics of the data source as well as algorithmic factors to estimate the square dimensions surrounding any point over which land cover will influence the metric. It is used later in the predictor selection process to simultaneously optimize model error rate and effective spatial resolution (see **Section 2.7.4**).

2.6.1. Aerial imagery metrics

A variety of metrics representing spectral response and spatial patterning were calculated from the aerial imagery (**Table 5**). Two main types of metrics were produced. Reflectance metrics, produced from the shadow-corrected 2015 image, are based on responses in different spectral bands from a single imaged pixel. Texture metrics, produced from the shadow-corrected multi-year merged image, are based on local variability in spectral responses, measured across a moving window incorporating numerous pixels. The processing is described in greater detail below.

Table 5. Aerial imagery-based predictive metrics the effective spatial resolution at which they respond, and a brief description or reference to a methodology.

Metric name(s)	Effective res (m)	Description
r1_md, r1_mx	3	Median and maximum red band value over source pixels
g1_md, g1_mx	3	Median and maximum green band value over source pixels
n1_md, n1_mx	3	Median and maximum near-IR band value over source pixels
u1_md, u1_mx	3	Median and maximum near-IR:green band contrast over source pixels
v1_md, v1_mx	3	Median and maximum near-IR:red band contrast (NDVI, Rouse et al. 1974, Tucker and Sellers 1986) over source pixels
w1_md, w1_mx	3	Median and maximum red:green band contrast over source pixels
x1_md, x1_mx	3	Median and maximum green:blue band contrast over source pixels
y1_md, y1_mx	3	Median and maximum red:blue band contrast over source pixels
r1a, r1b, r1c	3	Texture metric via filter ‘a’, ‘b’, ‘c’ at 1m resolution on red band
r2a, r2b, r2c	6	Texture metric via filter ‘a’, ‘b’, ‘c’ at 2m resolution on red band
r3a, r3b, r3c	9	Texture metric via filter ‘a’, ‘b’, ‘c’ at 3m resolution on red band
r4a, r4b, r4c	12	Texture metric via filter ‘a’, ‘b’, ‘c’ at 4m resolution on red band
r6a, r6b, r6c	18	Texture metric via filter ‘a’, ‘b’, ‘c’ at 6m resolution on red band
r9a, r9b, r9c	27	Texture metric via filter ‘a’, ‘b’, ‘c’ at 9m resolution on red band
rca, rcb, rcc	36	Texture metric via filter ‘a’, ‘b’, ‘c’ at 12m resolution on red band
rda, rdb, rdc	54	Texture metric via filter ‘a’, ‘b’, ‘c’ at 18m resolution on red band
rea, reb, rec	81	Texture metric via filter ‘a’, ‘b’, ‘c’ at 27m resolution on red band
rfa, rfb, rfc	108	Texture metric via filter ‘a’, ‘b’, ‘c’ at 36m resolution on red band
gRF, nRF, uRF, vRF, wRF	3–108	All the above combinations of resolution (R) and convolution filter (F) applied to green band, near-infrared band, near-IR:green contrast, near-IR:red contrast, and green:red contrast

Table 5 (continued). Aerial imagery-based predictive metrics, the effective spatial resolution at which they respond, and a brief description or reference to a methodology.

Metric name(s)	Effective res (m)	Description
ra_13, rb_13, rc_13	3	NDTI of *r1a* contrasted with *r3a*, *r1b* with *r3b*, and *r1c* with *r3c*
ra_26, rb_26, rc_26	6	NDTI of *r2a* contrasted with *r6a*, *r2b* with *r6b*, and *r2c* with *r6c*
ra_39, rb_39, rc_39	9	NDTI of *r3a* contrasted with *r9a*, *r3b* with *r9b*, and *r3c* with *r9c*
ra_4c, rb_4c, rc_4c	12	NDTI of *r4a* contrasted with *rca*, *r4b* with *rcb*, and *r4c* with *rcc*
ra_6d, rb_6d, rc_6d	18	NDTI of *r6a* contrasted with *rda*, *r6b* with *rdb*, and *r6c* with *rdc*
ra_9e, rb_9e, rc_9e	27	NDTI of *r9a* contrasted with *rea*, *r9b* with *reb*, and *r9c* with *rec*
ra_cf, rb_cf, rc_cf	36	NDTI of *rca* contrasted with *rfa*, *rcb* with *rfb*, and *rcc* with *rfc*
gF_RS, nF_RS, uF_RS, vF_RS, wF_RS	3–36	All the above combinations of convolution filter (F) and two resolutions (R,S) applied to green band, near-infrared band, near-IR:green contrast, near-IR:red contrast, and green:red contrast
d1c, d2c, d3c, d4c, d6c, d9c, dcc, ddc, dec, dfc	3–108	Cross-band contrast between *v1c & r1c, v2c & r2c, v3c & r3c, v4c & r4c, v6c & r6c, v9c & r9c, vcc & rcc, vdc & rdc, vec & rec, vfc & rfc*
e1c, e2c, e3c, e4c, e6c, e9c, ecc, edc, eec, efc	3–108	Cross-band contrast between *n1c & r1c, n2c & r2c, n3c & r3c, n4c & r4c, n6c & r6c, n9c & r9c, ncc & rcc, ndc & rdc, nec & rec, nfc & rfc*

Spectral metrics

Response metrics were produced for the red, green and near-infrared bands of the 1-meter resolution imagery. In addition, several vegetation indices were calculated from the raw band values: the normalized difference vegetation index (NDVI; Rouse et al. 1974, Tucker and Sellers 1986), and parallel contrast metrics between the near-infrared and green bands, and between the green and red bands. Each metric was summarized to the 3-meter mapping resolution by taking the median and the maximum 1-meter value within each 3-meter modeling pixel.

Nested texture metrics

Most information in high-resolution imagery is contextual and expressed in the spatial patterning of pixel neighborhoods; the eye's ability to identify many features based solely on the patterning and arrangement of gray-scale brightness values illustrates this point. We devised a method called *nested texture metrics* (NTM) to extract this information and provide it as predictor data to the modeling process. The texture metrics represent local variability at a range of pixel resolutions corresponding to distinct spatial scales at which various vegetation and landscape features occur.

Each of the spectral metrics described above was first median-aggregated[43] to a variety of coarser resolutions (2, 3, 4, 6, 9, 12, 18, 27, and 36 meters). We then used three different 3x3-cell[44] convolution filters to extract different aspects of patterning from each of the aggregated datasets as well as the original 1-meter dataset: (a) standard deviation of the center cell and the eight nearest neighbor (8NN) cells; (b) a 'speckle' filter, the absolute value difference between the center cell and the median of the 8NN, divided by the median of the 8NN and then smoothed by an additional 3x3-cell median filter; and (c) a non-trending variance filter accomplished via an alternating-cell convolution kernel $\begin{bmatrix} +4 & -5 & +4 \\ -5 & +4 & -5 \\ +4 & -5 & +4 \end{bmatrix}$. The results were converted to 3-meter resolution by a combination of median aggregation and cubic convolution resampling designed to maintain high-resolution detail.

Normalized difference texture index

We developed the normalized difference texture index (NDTI) to minimize the impact of variability in view angle and illumination characteristics between flight lines. The index works on the principle that because these artifacts affect textures similarly across a range of pixel resolutions, they can be partially canceled out by contrasting textures computed at two different pixel resolutions. Texture differences remaining after this cancellation result from image patterns at spatial scales intermediate between the two resolutions. NDTI metrics were produced by contrasting metrics computed at the following pairs of resolutions: 1m/3m, 2m/6m, 3m/9m, 4m/12m, 6m/18m, 9m/27m, and 12m/36m.

$$NDTI_{ab} = \frac{(\sigma_a - \sigma_b)}{(\sigma_a + \sigma_b)}$$

where a and b represent the two source texture resolutions and σ represents the source texture metric computed at the given resolution.

Cross-band contrast metrics

We produced another set of metrics to contrast corresponding metrics computed on NDVI against the red band, and on the near-infrared band against the red band. A formula like that used for NDTI was used, based only on the results from the 'c' convolution filter.

2.6.2. Satellite imagery metrics

We calculated a variety of metrics based on mathematical transformations of the satellite imagery (**Table 6**). The metrics differed somewhat depending on whether the source imagery was obtained by Sentinel-2, Landsat-8 or Landsat-5. All applicable metrics were produced for the current midsummer

[43] GIS data is typically aggregated to a coarser resolution by taking the mean value of the finer resolution input pixels across each of the output pixels. Summarizing by the median value instead reduces smoothing near land cover transitions and increases the isolation of scale-dependent texture signals.

[44] The term *cell* is generally synonymous with *pixel*, but we mean it to convey a more abstract conception—generally in the context of a data-processing algorithm—than that conveyed by *pixel*, which is usually associated with the local contribution to some larger "picture."

and minimum-snow images and for the historic midsummer image, but only those from current images were used in modeling.

Table 6. Satellite imagery-based predictive metrics the effective spatial resolution at which they respond, and a brief description or methodology reference.[A] Where two resolutions are shown, the first is for Sentinel imagery, the second for Landsat.

Metric name(s)	Effective res (m)	Description
grn	10, 30	Green reflectance: 543–577 nm (Sentinel-2), 530–590 nm (Landsat-8) or 520–600 nm (Landsat-5)
red	10, 30	Red reflectance: 650–680 nm (Sentinel-2), 640–670 nm (Landsat-8) or 630–690 nm (Landsat-5)
re1	20	Red edge reflectance: 698–712 nm (Sentinel-2 only)
re2	20	Red edge reflectance: 733–747 nm (Sentinel-2 only)
re3	20	Red edge reflectance: 773–793 nm (Sentinel-2 only)
nir	10, 30	Near-infrared reflectance: 785–899 nm (Sentinel-2), 850–880 nm (Landsat-8) or 760–900 nm (Landsat-5)
sw1	20, 30	Shortwave reflectance: 1565–1655 nm (Sentinel-2), 1570–1650 nm (Landsat-8) or 1550–1750 nm (Landsat-5)
sw2	20, 30	Shortwave reflectance: 2100–2280 nm (Sentinel-2), 2110–2290 nm (Landsat-8) or 2080–2350 nm (Landsat-5)
temp	100	Thermal band response: 10.60–11.19 µm (Landsat-8 only)
ndvi, ndvip	10, 30	Normalized difference vegetation index (Tucker and Sellers 1986)
ndmi, ndmip	20, 30	Normalized difference moisture index (Wilson and Sader 2002)
ndfi, ndfip	20, 30	Normalized difference forest index = *ndvi* + *ndmi*
nbr, nbrp	20, 30	Normalized burn ratio (Key and Benson) 2002
ndsi, ndsip	20, 30	Normalized difference snow index (Hall et al. 1995)
ndgr, ndgrp	10, 30	Normalized contrast between *grn* and *red*
ndng, ndngp	10, 30	Normalized contrast between *nir* and *grn*
ndsw, ndswp	20, 30	Normalized contrast between *sw1* and *sw2*
tcb	20, 30	Tasseled cap brightness (Kauth and Thomas 1986, Huang et al. 2002)
tcg	20, 30	Tasseled cap greenness (Kauth and Thomas 1986, Huang et al. 2002)
tcw	20, 30	Tasseled cap wetness (Kauth and Thomas 1986, Huang et al. 2002)
di	20, 30	Disturbance index (Healey et al. 2005)
ndre, ndrep	20	Normalized difference red edge index (Barnes et al. 2000, Sentinel-2 only)
ccci	20	Canopy chlorophyll content index (Barnes et al. 2000, Sentinel-2 only)
mcari	20	Modified chlorophyll absorption ratio index (Daughtry et al. 2000, Sentinel-2 only)
resav	20	Red edge soil-adjusted vegetation index (Cao et al. 2013, Sentinel-2 only)

[A] The indices ending in 'p' were developed during this work. They were calculated by adding 2 to the denominator of the standard formula for the metric, to compensate for index overestimation on dark surfaces such as water and deep shadow.

2.6.3. Topographic metrics

A variety of metrics describing the influence of local topography on vegetation composition were calculated (**Table 7**). These were derived primarily from the lidar bare earth elevation dataset, except in areas beyond that dataset's extent, where 10-meter resolution 3DEP data were substituted. The more complex novel metrics created during this project are briefly described here.

Table 7. Topographic predictive metrics the effective spatial resolution at which they respond, and a brief description or reference to a methodology.

Metric name(s)	Effective res (m)	Description
elev	3	Bare earth elevation.
slope	3	Slope in degrees (Esri 2013).
east, south	3	"Eastiness" = sin(aspect) and "southiness" = sin(aspect-90°).
cur6, cur30, cur150, cur750	6, 30, 150, 750	3x3-cell total curvature (Esri 2013) from elevation aggregated to 6m, 30m, 150m and 750m resolution.
cpl6, cpl30, cpl150, cpl750	6, 30, 150, 750	3x3-cell planimetric curvature (Esri 2013) from elevation aggregated to 6m, 30m, 150m, and 750m resolution.
cpr6, cpr30, cpr150, cpr750	6, 30, 150, 750	3x3-cell profile curvature (Esri 2013) from elevation aggregated to 6m, 30m, 150m and 750m resolution.
heat	3	Relative heat load (McCune and Keon 2002).
raddir, raddur	30	Direct solar radiation and duration of direct illumination across full year (Esri 2013); distinct from heat load in that cast topographic shadows are modeled.
topodry	30	Elevation-scaled heat index = *raddir* * (1 – (*elev* / highest elev in WA)).
mp126, mp630, mp3150	6, 30, 150	Morphometric protection (SAGA-GIS, Conrad et al. 2015) from elevation aggregated to 6m over 126m radius, to 30m over 630m radius, and to 150m over 3150m radius.
tpp60, tpp300, tpp1500, tpp7500	6, 30, 150, 750	Topographic position percentile, the percentile rank of cell elevation relative to surrounding elevations within a 60m, 300m, 1500m and 7500m radius.
tpmi60, tpmi300, tpmi1500, tpmi7500	6, 30, 150, 750	Minimum elevation differential within 60m, 300m, 1500m and 7500m. See text for methodology.
tpma60, tpma300, tpma1500, tpma7500	6, 30, 150, 750	Maximum elevation differential within 60m, 300m, 1500m and 7500m. See text for methodology.
cold60, cold300, cold1500, cold7500	6, 30, 150, 750	Cold air accumulation calculated over surrounding 60m, 300m, 1500m and 7500m. See text for methodology.
rough9, rough30, rough90, rough270	3, 10, 30, 90	Surface roughness at 9m, 30m, 90m and 270m scales. See text for methodology.

Minimum and maximum elevation differentials

We devised two multi-resolution metrics to quantitatively represent landform position. Four bisecting lines of length 21 times the cell resolution were constructed for each cell, oriented in the N-S, NE-SW, E-W, and SE-NW directions. The mean elevation along each line was determined, and differences between the central cell's elevation and each of the four means were calculated. The minimum of these four differences (*minimum elevation differential* or *tpmi*) and the maximum

(*maximum elevation differential* or *tpma*) are relevant with respect to landform position. For example, a peak would have high values of both *tpma* and *tpmi*, while a level ridgeline would have a high *tpma* and a *tpmi* near zero. A gap in a ridgeline would have a high *tpma* and a fairly large negative *tpmi*. The metrics were calculated at a variety of cell sizes to represent terrain morphology at a variety of spatial scales.

Cold air accumulation

Cold air accumulation in basins is a major driver of vegetation patterns in mountainous terrain. We developed an original approach for simulating this process, using the four elevation differentials created above. Locations at which the sum of the elevation differentials across perpendicular axes is a negative number have some tendency to accumulate cold air draining from above. The greater the magnitude of this negative number, the greater will be the tendency for cold air to enter from above and become trapped, and the colder that air is likely to be.

For the four elevation differentials ed_{NS}, ed_{EW}, ed_{NESW} and ed_{SENW} we found the minimum sum of each of the perpendicular pairs:

$$ed_{\perp,min} = \min(ed_{NS} + ed_{EW}\ ,\ ed_{NESW} + ed_{SENW})$$

By analogy with the compound topographic index (Moore 1991)—a hydrologic metric that similarly integrates the influence of a size-varying contribution area with the local tendency to disperse that input—we represented cold air accumulation at a cell using:

$$cold = \ln\left(\frac{ed_{\perp,min}}{s}\right)$$

where s is the slope in the downward direction from the cell at the same spatial scale over which the elevation differentials were calculated. Cold air accumulation was determined at each of the cell sizes for which elevation differentials were produced.

Surface roughness

We defined surface roughness as local variability in aspect that is non-trending across an analysis window, scaled up by the local slope. The non-trending criterion is important—for example, a window centered on a north-south oriented ridgeline would show a strong change in aspect from west-facing to east-facing, but this would not indicate surface roughness. To accomplish this, we again used the alternating-cell convolution kernel $\begin{bmatrix} +4 & -5 & +4 \\ -5 & +4 & -5 \\ +4 & -5 & +4 \end{bmatrix}$, this time applied to four transformations of aspect: $\sin(asp)$, $\sin(asp - 45°)$, $\sin(asp - 90°)$ and $\sin(asp - 135°)$, summed these four directional measures of aspect variability, and multiplied by the mean slope across the analysis window. Roughness was computed at a range of spatial scales.

2.6.4. Hydrologic metrics

The hydrologic metrics were derived from processing within a landscape context rather than from a simple pixel-based perspective, since they depend on upstream areas in addition to the immediate surroundings. We first created a hydrologic flow accumulation layer based on the bare earth

elevation, correcting for poorly modeled flow due to lack of information on road culvert locations. We used the flow accumulation layer to create a channel network, calibrating it using an NPS streams data layer. The channel network was used as an input to a variety of distance metrics describing proximity to channels exceeding various flow thresholds. The predictive hydrologic metrics are shown in **Table 8**.

Table 8. Hydrologic predictive metrics the effective spatial resolution at which they respond, and a brief description or reference to a methodology.

Metric name(s)	Effective res (m)	Description
vd_drain, vd_perm, vd_major	6	Vertical distance above slope concavities, permanent channels and major river networks (Conrad et al. 2015)
hd_drain, hd_perm, hd_major	6	Horizontal distance to slope concavities, permanent channels and major river networks
dtw	6	Cartographic depth to water index (White et al. 2012)
wetness	6	SAGA wetness index (Conrad et al. 2015), closely related to Compound Topographic Index (Moore 1991)
upland	6	Log-scaled cost distance to channel network, see text

Flow accumulation and channel networks

Hydrologic *flow accumulation* is a spatial representation of the catchment area contributing to flow at each gridded location in a drainage network. Its computation was important both as a step in the channel delineation process and also as a key input needed to generate several predictive metrics. The flow accumulation algorithm in SAGA-GIS (Conrad et al. 2015), when used to delineate channel networks, produced anastomosing effects in flat areas and appeared to realistically represent hydrologic processes for incorporation into predictive metrics.

Although lidar data allows flow modeling with much greater spatial detail than a 3DEP DEM would permit, this detail can be a source of difficulties as well. For example, where culverts allow water to pass under roads and remain in its natural path, modeled flow paths may be blocked and diverted by road prisms. We reduced such problems and also eased computation by performing hydrological modeling at 6-meter resolution rather than at the full lidar resolution. We reduced the resolution by aggregating based on the minimum, setting each 6-meter cell equal to the lowest value of the 4 constituent 3-meter cells. This resulted in a reduced number of obstructed flow paths in subsequent modeling, as compared to aggregating based on the mean. In order to represent the impact of spatial precipitation patterns on channel development, we created a weighted grid by rescaling PRISM annual precipitation to a fraction of the maximum value in the study area. We then filled sinks in the elevation grid, using the Wang & Liu (2006) method with *minslope* = 0.01, and modeled flow accumulation based on the weighted precipitation grid, using SAGA's **Catchment Area (Top-Down)** method with multiple flow directions and *convergence* = 1.1.

Despite the minimum aggregation step above, flow paths along the upslope sides of roads continued to be a problem. Mismodeled flow eventually will find its way across the road, resulting in a stream in an incorrect location. A roads layer with very high spatial accuracy was needed in order to allow a targeted approach for flow correction near roads without compromising the quality of flow paths in other areas. No existing roads layers were sufficient, so we hand-digitized park roads based on the lidar slope image. The centers of all existing roads were digitized; old roads were also digitized if the prism was still apparent in the slope image. For all areas adjacent to roads, the 6-meter resolution elevation was set to the minimum elevation in the surrounding 3x3-cell window, in an attempt to route flow downslope across the roads. The sink filling and flow accumulation procedures above were then repeated.

We used the flow accumulation results to delineate channel networks, also in SAGA-GIS. Three alternate channel networks were created from the flow accumulation result. One was calibrated to represent visible slope concavities, another to represent permanent channels, and the third to represent only unconstrained rivers in major valleys. Various minimum thresholds of flow necessary to result in a channel were tested; the resulting networks were visually compared to stream representations in USGS 1:24,000 quad sheets. The best match to the represented permanent streams was found using a flow accumulation threshold of 25,000, corresponding to an average catchment area of about 90 hectares. A threshold of 10,000,000 (corresponding to about 36,000 ha) was used for major rivers; this resulted in delineation of channels downstream of the approximate location where their floodplains begin to widen substantially. We chose a threshold of 1,000 (corresponding to about 3.6 ha) for slope concavities, because it produced results that generally matched the representation of discernible hollows on the quad sheets.

Riparian influence and metrics generation

We devised a metric to express the degree of floristic riparian influence at any location. The first step was to determine the total flow quantity associated with each section of the channel network. The channel network was broken into discrete channel reaches defined by network intersections. Many channel segments were composed of anastomosing flow pathways, in which flow was modeled in several adjacent parallel paths; it was therefore necessary to consider the several paths as all contributing to a single total flow value. We accomplished this by associating each flow accumulation cell with the nearest delineated reach[45] and averaging across reach length.

We classified channel reaches into five categories based on average reach flow, with thresholds between the categories spaced in a regular geometric progression ranging from the minimum to the maximum channel reach flow in the study area. We then created a cost function to describe the degree of riparian influence in the perpendicular direction away from the channel. The cost function was proportional to the square of slope, which emphasized slope breaks and was able to represent physiographic features such as fluvial terraces and natural levees. We calculated the least cost distance from each cell to each of the five channel size categories using this function.

[45] "Distance" to the reach was evaluated via a cumulative slope cost function.

The riparian influence metric was fit to its practical impact on species composition by examining the cost function values at the locations of training plots assigned to riparian vs. non-riparian associations. This resulted in an estimate of a cost function cutoff for each of the five flow categories that most accurately separated the plots with riparian floristics from those with upland floristics. A logarithmic relationship was found to best fit the relationship between the five cost function cutoffs and the mean flow quantity across all reaches in each of the five flow categories. We then iteratively modified the initial cutoffs until they exactly fit the logarithmic model. For each of the flow categories, we assumed that no further floristic riparian influence would be exerted beyond the cost distance cutoff. Finally, an "uplandness" index was created using:

$$upland = \log_{10}\left[1 + \min\left(\frac{C_A}{T_A}, \frac{C_B}{T_B}, \frac{C_C}{T_C}, \frac{C_D}{T_D}, \frac{C_E}{T_E}\right)\right]$$

where $C_{A..E}$ represent the slope-based cost distances to each of the five flow categories and $T_{A..E}$ represent the cost cutoffs used to define the extent of riparian influence for each category.

2.6.5. Climate metrics

The climate data required no additional processing to form predictive metrics. The predictors (**Table 9**) were simply the 1981–2010 normals provided by the PRISM Climate Group (2019).

Table 9. Climate predictive metrics the effective spatial resolution at which they respond, and a brief description or reference to a methodology.

Metric name(s)	Effective res (m)	Description
ppt_jan, ppt_apr, ppt_jul, ppt_oct	~800	Average precipitation for month.
tmax_jan, tmax_apr, tmax_jul, tmax_oct	~800	Average daily maximum temperature for month.
tmin_jan, tmin_apr, tmin_jul, tmin_oct	~800	Average daily minimum temperature for month.
tdew_jan, tdew_apr, tdew_jul, tdew_oct	~800	Average daily mean dew point temperature for month.
vmax_jan, vmax_apr, vmax_jul, vmax_oct	~800	Average daily maximum vapor pressure deficit for month.

2.6.6. Vegetation canopy metrics

In addition to the 3-meter resolution artifact-filtered vegetation height, a variety of predictors were formed by summarizing canopy height information over moving windows. **Table 10** provides an overview of the metrics representing the vegetation canopy; they are described in greater detail below.

Table 10. Vegetation canopy predictive metrics the effective spatial resolution at which they respond, and a brief description or reference to a methodology.

Metric name(s)	Effective res (m)	Description
ht3m	3	Local vegetation height
ht50, htmax, htsd	27	Median height, maximum height, and standard deviation of height within 15m radius circular moving window
htt50, htw50	27	Median height value of heights of tree-sized (over 5m) and woody-sized (over 80cm) vegetation, within 15m radius circular moving window
cctree, ccwood	27	Fractional cover of tree-sized and woody-sized vegetation, within 15m radius circular moving window
errmn	27	Canopy elevation relief ratio (see text) within 15m radius circular moving window
h1b3b	3	NDTI of 1m 'b' texture metric contrasted with 3m 'b' texture metric (on height)
h2b6b	6	NDTI of 2m 'b' texture metric contrasted with 6m 'b' texture metric (on height)
h3b9b	9	NDTI of 3m 'b' texture metric contrasted with 9m 'b' texture metric (on height)
h4bcb	12	NDTI of 4m 'b' texture metric contrasted with 12m 'b' texture metric (on height)
h6bdb	18	NDTI of 6m 'b' texture metric contrasted with 18m 'b' texture metric (on height)
h9beb	27	NDTI of 9m 'b' texture metric contrasted with 27m 'b' texture metric (on height)
hcbfb	36	NDTI of 12m 'b' texture metric contrasted with 36m 'b' texture metric (on height)

Several summary metrics were generated over a 15-meter radius circular window. To minimize artifacts near vegetation transitions, most were based on the median rather than the mean. The predictors *ht50*, *htmax*, and *htsd* were calculated respectively as the median value, maximum value, and standard deviation of the heights within the circular window.

We created several predictors to describe the height and density of the upper vegetation layers, if those layers were present within the circular window. The median tree height *htt50* was calculated as the median value of the height values that exceeded the tree height threshold of five meters. Rather than indicating the height of the dominant vegetation, *htt50* describes the height of any tree layer present, no matter how sparse it is. The median woody vegetation height *htw50* was defined similarly, but using a threshold of 80 centimeters, assuming that most height values over that correspond to woody vegetation of some kind. Tree and woody vegetation canopy cover (*cctree* and *ccwood*) were defined as the fraction of the window with height values over five meters and 80 centimeters, respectively.

Another descriptor of vertical canopy structure was derived based on the elevation-relief ratio of Pike and Wilson (1971). The mean canopy elevation relief ratio *errmn* specifies the fractional distance of the mean canopy height *htmean* between the minimum *htmin* (usually zero) and maximum height *htmax* in the focal window. It was computed by:

$$errmn = \frac{htmean - htmin}{htmax - htmin}$$

The NDTI was computed based on the ‘b’ texture metric from the height raster, at the same range of spatial scales as described in **Section 2.6.1**. These metrics represent measures of vertical canopy roughness at different horizontal scales. Finer scale NDTI metrics are sensitive to individual trees and small canopy openings, while coarser scale metrics are more responsive to variability in overall canopy height and larger openings. Several other roughness metrics were investigated, such as aerodynamic roughness (e.g., Menenti and Ritchie 1994) and fractal dimension (e.g., Isaaks and Srivastava 1988), but the gain in information expected from these metrics appeared to not justify the cost and complexity of producing them across the project area.

The availability of highest hit lidar data opened up another dimension of predictor data for modeling at OLYM, providing high-resolution information that is not available from other sources. **Figure 11** illustrates the sort of clarity and detail that it provided for modeling and mapping forests.

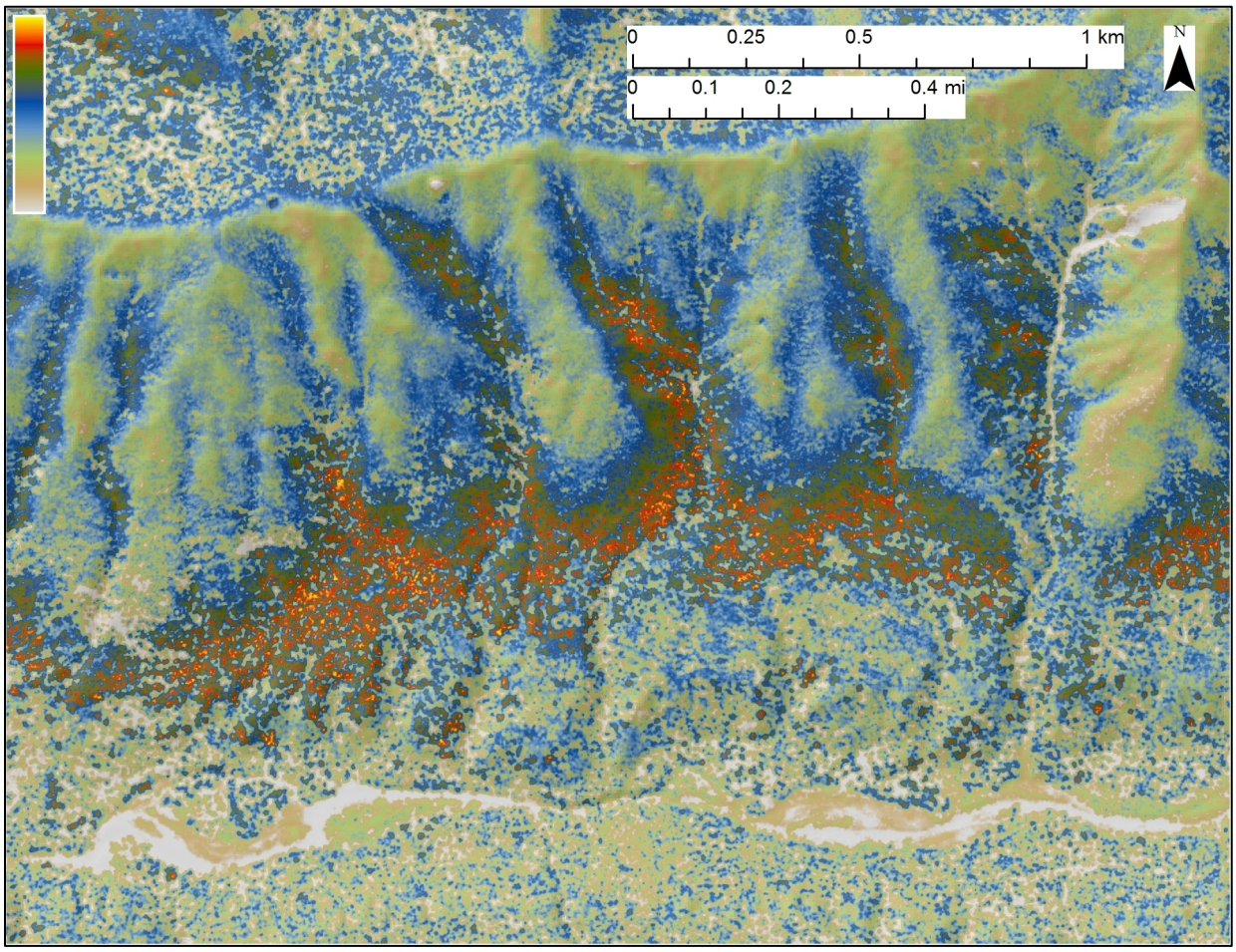


Figure 11. Lidar-derived vegetation height in the Bogachiel Valleyon the south side of Sugarloaf Mountain. Some of the park’s big trees are found on the bench and toe slope in the center of the image. The color ramp used for increasing height is shown in the upper left corner; the data may be uncalibrated so no corresponding figures are reported. The impressive old growth here mapped mostly as C04–MOIST WESTERN HEMLOCK, DOUGLAS-FIR AND FOAMFLOWER FOREST.

2.7. Modeling

We used a machine learning algorithm, *random forests* (RF; Breiman 2001, Liaw and Wiener 2002), to build models for predicting map class presence using the quality-controlled training plots resulting from the work in **Section 2.3**. We used RF because of its tendency to avoid overfitting to training data and its ability to isolate signals in noisy datasets (Cutler et al. 2007). The large number of map classes, with widely varying quantities of available training data, presented a modeling challenge: how to simultaneously produce models that are good at both "easy" prediction tasks (e.g., discriminating between low and high elevation types) and "hard" tasks (e.g., discriminating between two tall shrubland types occurring in similar settings), while avoiding bias against the rarest classes and also making maximum use of all available training data. To address this, we wrapped the RF algorithm in a factorial binary process in which each map class was modeled against every other. This allowed each model to specialize in distinguishing a single pair of map classes, choosing appropriate predictors for that task. During the prediction phase each class "competed" with each other class; the class with the lowest cumulative loss margin across all contests at a pixel was considered the best answer there. The predictor selection, model creation, and model prediction phases discussed below all ran on binary models.

The vegetation canopy predictors described in **Section 2.6.6** could only be produced where lidar data were available. This forced us to produce a separate set of models in order to map the coastal strip and other sections of the project area that lacked lidar. The steps described in **Sections 2.7.3–6** were performed twice: the *lidar run* included the lidar predictors, and was used wherever lidar data were available and appeared reliable; the *no-lidar run* left them out, and was used only where lidar data were unavailable or unreliable.

2.7.1. Model predictor data

All the metrics discussed in **Section 2.6** were resampled to a fixed 3-meter resolution grid over the coincident extent of all metrics. The resampling method used depended on the data source. We used nearest neighbor resampling to maintain the finest resolution possible for all metrics derived from NAIP and lidar; the predictor sampling grid was taken from these rasters to prevent any spatial shifting. Satellite imagery was resampled using cubic convolution, which results in less smoothing than bilinear interpolation and maintains crisper boundaries. The non-imagery layers were resampled using bilinear interpolation.

2.7.2. Model training data

Following the quality control process, the training plots represented relatively continuous patches of the assigned map class, spanning the full area defined by the plot center location and radius. Any patches of alternate types within the plot were assumed to be less than nine meters on a side. Training data were created from the predictor metrics by extracting the 3-meter pixel values at 13 points distributed across each training circle, with the most distant four points lying on the circumference

(**Figure 12**).[46] The primary reason for extracting data from multiple locations at each plot was the necessity of training models at the same spatial scale at which they were predicted.[47] In addition, this scheme allowed better representation of the range of predictor variation within each plot, including providing training data near transitions to adjacent vegetation types. The assigned association and map class calls and the extracted predictor values were then imported into R using functions provided by the **rgdal** package (Bivand et al. 2014). The training data represented 51 distinct map class calls; flowing water and impounded water were modeled separately, but were later merged into the single map class W81–FRESH WATER.

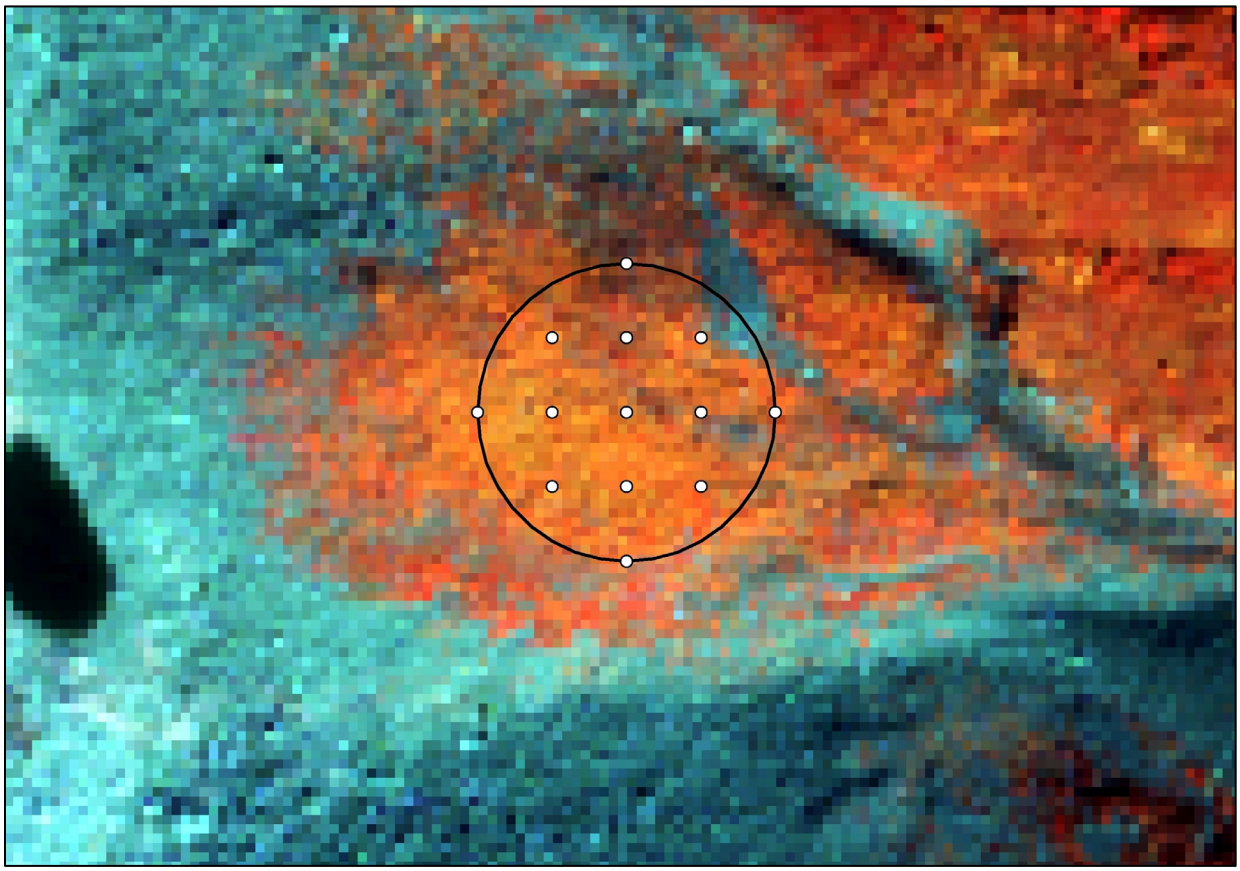


Figure 12. Training data extraction Predictors were extracted from the 3-meter resolution metrics at 13 points distributed across each training circle. The vegetation patch represented was assigned to H53–SHOWY SEDGE AND SITKA VALERIAN MEADOW; the imagery is color-infrared 2015 NAIP at 1-meter resolution.

[46] Because the 13 component samples from each plot are not statistically independent, we used only one of the 13 in any given random forests tree during the predictor selection and model creation steps. This avoided introducing pseudo-replication of training data and preserved the independence of the out-of-bag samples, while making use of the predictor variability within each plot.

[47] A commonly used alternative, summarizing predictor data over the plot area, would have introduced artificial smoothing into the training data that would not be applicable for predicting onto finer resolution pixels.

2.7.3. Model binarization

Each two-class combination of the 51 modeled map classes was treated separately, resulting in 1,275 distinct binary models.[48] This allowed each model to specialize in a single task—distinguishing two classes from one another—and gave us the freedom to treat issues of predictor collinearity more sensitively. For example, over the geography defined by all training samples in NOCA, there is a very strong negative correlation between elevation and maximum January temperature. But within the environmental subspace defined by the training plots assigned to C05–WESTERN HEMLOCK, DOUGLAS-FIR AND SWORD FERN FOREST and C06–WESTERN HEMLOCK, DOUGLAS-FIR AND SALAL FOREST, those variables are only weakly correlated, with maximum January temperature being a very strong predictor and elevation comparatively weak. Using both in a model based on training data throughout NOCA would violate standards against excessively correlated predictors. But there is no such violation for the single model C05 versus C06 and to exclude either predictor on this basis would unnecessarily reduce the model accuracy.

2.7.4. Predictor selection

We developed a novel predictor selection method to use with our multi-resolution predictor datasets, which reduced predictor collinearity[49] while also optimizing model accuracy, model effective spatial resolution, and the efficiency of the prediction process.

Initial selection

We used a stepwise variable selection process coded in R, which was based on maximizing RF cross-validated model accuracy at each step.[50] We organized the predictors into ten tiers based on the effective spatial resolution at which they were calculated,[51] with the finest scale predictors—the 3-meter resolution NAIP band responses and lidar-derived canopy height—in the first tier.

At each tier, the process cycled through all available predictors, building 100 forests of 501 trees each, with each forest built from a single randomly selected point of the 13 for each plot. For each of the two map classes in the model, the *out-of-bag error rate*[52] for each plot, e_p, was compiled over

---

[48] We handled the computationally intensive process of predictor selection, model creation and model prediction at 3-meter resolution using three standard desktop computers, each running between three and five instances of R or Python simultaneously. They were connected to a network-attached storage device that hosted the training data, predictor data, and a shared status file that allowed the processes to distribute the tasks amongst themselves.

[49] Inclusion of substantially correlated predictors causes RF to overfit to those predictors, which is a major concern because our training data were gathered from such a small fraction of the project area.

[50] We considered using a process guided by an importance measure returned by RF, as in Evans and Cushman (2009). However, we found that a predictor's contribution to model accuracy is strongly dependent on which other predictors are included, and that an importance measure returned from a model based on all predictors was not indicative of its potential utility in a model based on a small subset.

[51] However, despite the availability of high-resolution topography from lidar across much of the park, we considered topographic and hydrologic predictors together with mid-resolution satellite imagery—after Sentinel imagery but before Landsat—to keep the emphasis on existing conditions as opposed to environmental setting.

[52] RF generates this by testing each tree of the model against the samples that were withheld from creating it.

each of the forests and converted to an estimate of the probability of plot misclassification by a single forest.[53] This probability estimate was then averaged across all plots to produce an overall error rate estimate for the model including the newly introduced predictor.[54] The predictor in the tier that resulted in the greatest decrease in model error rate was selected; any predictors (in that tier or others) with an absolute-valued Spearman rank correlation of 0.8 or greater to the selected predictor were eliminated from further consideration. If no predictors within the tier resulted in a decrease in model error rate, consideration moved to the following tier. After a predictor was selected, consideration always moved to the first tier again.[55]

Climate variables can act as proxies for geographic location, as they are generally arranged along broad spatial gradients. Their use as predictors can present a severe risk of overfitting to training data whose collection has been determined more by convenience than by a random sample. Because our climate predictors were derived from approximately 800-meter resolution data, they were in the final selection tier. We additionally limited models to only one climate predictor, to reduce the likelihood of overfitting to our often spatially constricted training data.

Often there was an inherent tradeoff between accuracy and spatial resolution. If satellite imagery or coarser scale texture metrics provide key information that is lacking in finer scale data, their use will increase accuracy but will also coarsen the model's resolution (**Figure 13**; from lidar run).

[53] The expected misclassification by a single forest was of interest because the map was made based on a single forest. This step assumed normal distribution of e_p across forests.

[54] The model error rate here is defined as the higher error rate of the two modeled classes. By optimizing this quantity, rather than the overall (average) model error rate, we kept the error rate of the two classes balanced, which was an important assumption made by our prediction method.

[55] Because a predictor's value may not be recognized until a compatible predictor has been included.

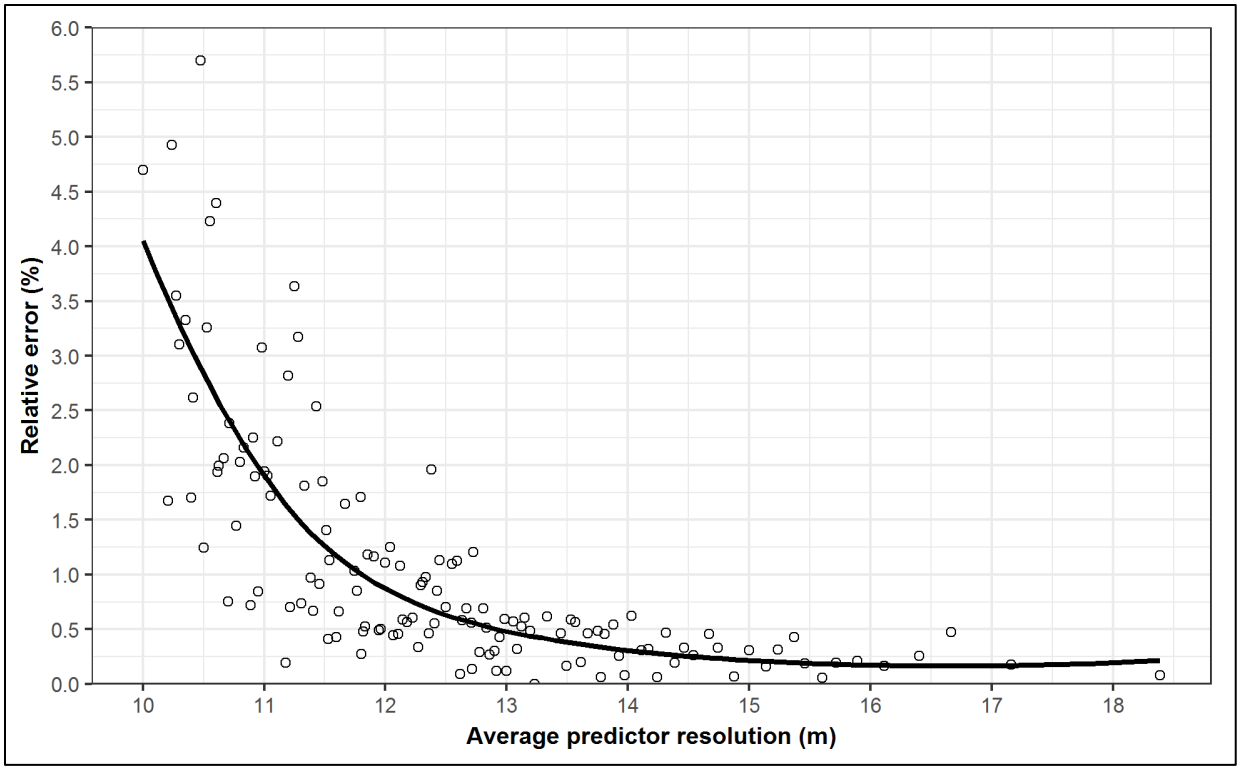


Figure 13. Average predictor resolution vs. median relative model error(the error increase attributable to leaving out coarser predictors). Predictors finer than 10-m resolution were treated as 10 m here; greater averages indicate increasing incorporation of coarser predictors in the model. The data were derived from only models with final error of 1% or greater, with a median final error of 3.6%; this would have been 7.7% using only predictors of 10-m or finer resolution. The best fit line is from a loess smoothing function.

Predictor switching

As seen in **Figure 13**, there is an optimal resolution at which to produce a model, which takes advantage of some of the predictive power of coarser resolution predictors, while maintaining responsiveness to fine scale vegetation transitions.[56] The predictor selection routine described above simply tried to minimize model error, but some of this may come at the unnecessary expense of coarser resolution. To address the tradeoff between the two, we created a new metric *errxres* that combined both model error and average predictor resolution:

$$errxres = avgres(err + 1\%)$$

where *err* is the model error (in percent) and *avgres* is the average predictor resolution. Starting with the predictors selected in the previous phase, we used another R script to drop the last selected predictors until the value of *errxres* was minimized. We then tested each of the remaining predictors, finding the model error rate that resulted from substituting any highly correlated predictors for them.

[56] The assumption is being made that the average spatial resolution of the predictors included in model is related to the effective resolution at which it "maps." Since RF is an inherently non-linear process, this is not necessarily true, though it is intuitively appealing.

Any substitutions that resulted in lowering *errxres* were accepted. The resulting predictor list was saved as an alternative set.

Choosing best set

To choose between the two sets of predictors produced, we used different decision-making criteria depending on whether we wanted to prioritize error rate or mapping resolution for the model. If the two map classes in the model were both larger patch size types (e.g., most conifer-dominated map classes), or if their environmental envelopes were so distinct that they wouldn't be found in close proximity to one another, we concluded that high spatial resolution in the resulting map was not as important as model error rate. In this case, we kept the set of predictors that resulted in the lowest error rate.

For pairs of map classes in which fine grain mapping was a high priority, we kept the predictor set that minimized the product $avgres(err')$, where *err'* was a transformed version of *err* that prioritized the reduction of model error to 5%, but only gave partial credit for reducing error lower than that.[57] We did this because the training samples were not perfectly pure[58] and we wanted to prioritize predictor resolution once a low error rate had been achieved. For example, a model to distinguish between a meadow and a woodland may have been trained with meadow samples that had occasional scattered trees. The best model in this case may have been one with a non-zero error rate against the training data. **Table 11** lists the most frequently included predictors across all binary models.

[57] $err' = \begin{cases} \max\left(4, \frac{(err+15)}{4}\right) & if\ err \leq 5 \\ err & if\ err \geq 5 \end{cases}$

[58] See **Figure 12**. Small patches where an alternative map class might be preferable are present in many training plots.

Table 11. Most frequently used predictorsin each selection tier (lidar run). Up to ten predictors are shown for each tier, provided they were used in at least 2% of the models.[A] The tables in **Section 2.6** provide descriptions for each predictor.

Tier	Resolution or type	Predictor names[B] and number of models in which used (in parentheses)
1	3 meters	*c_ht3m* (665), *n_d1c* (462), *n_e1c* (363), *n_g1b* (361), *n_w1c* (347), *n_g1c* (337), *n_n1c* (330), *r_x1_mx* (325), *n_v1c* (304), *n_w1a* (295)
2	6 meters	*c_h1b3b* (431), *n_wb_13* (235), *n_nb_13* (222), *n_wa_13* (212), *n_wc_13* (204), *n_ub_13* (201), *n_vc_13* (195), *n_nc_13* (193), *n_vb_13* (192), *n_d2c* (189)
3	9–10 meters	*s_nir* (252), *s_ndgr* (242), *s_grn* (204), *s_red* (170), *s_ndng* (138), *n_e3c* (110), *s_ndvi* (107), *n_d3c* (85), *n_g3b* (85), *n_w3c* (81)
4	12 meters	*c_h2b6b* (132), *n_ua_26* (65), *n_wa_26* (61), *n_va_26* (55), *n_nb_26* (51), *n_d4c* (49), *n_ra_26* (46), *n_wb_26* (46), *n_ub_26* (43), *n_vb_26* (43)
5	18–20 meters	*s_ndsw* (77), *s_sw1* (75), *s_tcw* (69), *s_sw2* (60), *c_h3b9b* (55), *s_di* (43), *s_nbr* (35), *s_tcb* (35), *s_ndmi* (30), *n_ua_39* (29)
6A	topographic	*t_elev* (399), *t_topodry* (122), *t_slope* (96), *t_tpp7500* (93), *t_tpmi7500* (88), *t_tpma7500* (85), *t_cpl750* (77), *t_raddir* (77), *t_cold7500* (74), *t_raddur* (62)
6B	hydrologic	*h_hd_major* (163), *h_vd_major* (143), *h_wetness* (139), *h_upland* (108), *h_dtw* (77), *h_vd_perm* (49), *h_hd_perm* (38), *h_vd_drain* (32), *h_hd_drain* (31)
7	27–30 meters	*c_htmax* (31)
10	climate	*p_vmax_jan* (37), *p_ppt_jul* (32), *p_clidry* (26), *p_vmax_jul* (26)

[A] The 'p' variants of the summer satellite imagery metrics ("s") were omitted from modeling due to a programming error.

[B] Predictor names are preceded by a letter indicating to which source group they belong: "c" indicates vegetation canopy, "e" is late summer minimum-snow satellite imagery, "h" is hydrologic, "n" is aerial imagery NTM, "p" is climate, "r" is aerial imagery reflectance, "s" is summer satellite imagery, and "t" is topographic.

Additional predictors for abiotic map classes

Many models for abiotic map classes had a small number of predictors selected, since the most obvious difference between these and vegetated types is their lack of vegetation, which is easily ascertained from NAIP imagery. While these very simple models worked well under normal circumstances, we found that in deep shadows these models often performed poorly. The abiotic classes are generally restricted to environments that are easily described in terms of topographic and hydrologic metrics. For example, impounded water is found in areas with low slope and high topographic wetness, and barren colluvial deposits are found in concave areas with positive curvature. We added appropriate predictors to models involving these map classes to make sure they remained restricted to reasonable locations.

2.7.5. Model creation

We built a random forest of 507 trees for each map class pair,[59] using the predictors selected above and specifying a sample size for each class equal to the minimum number of training plots available

[59] 39 trees were generated for each of the 13 sample points at each training plot.

for either class.[60] The resulting model was saved for use later in the prediction phase. We then estimated model error rates using 1000 bootstrap samples. Each was constructed by holding out one plot from the least common class and a proportional number from the most common class, again randomly selecting from the 13 sample points available at each plot for both training and test sets. **Figure 14** illustrates the cumulative probability across error rate for all binary models in the no-lidar run. Of the 1,275 models, 43% had an error rate of zero. Substantial error is concentrated in a fairly small number of models; 94% of the models showed less than 10% error.

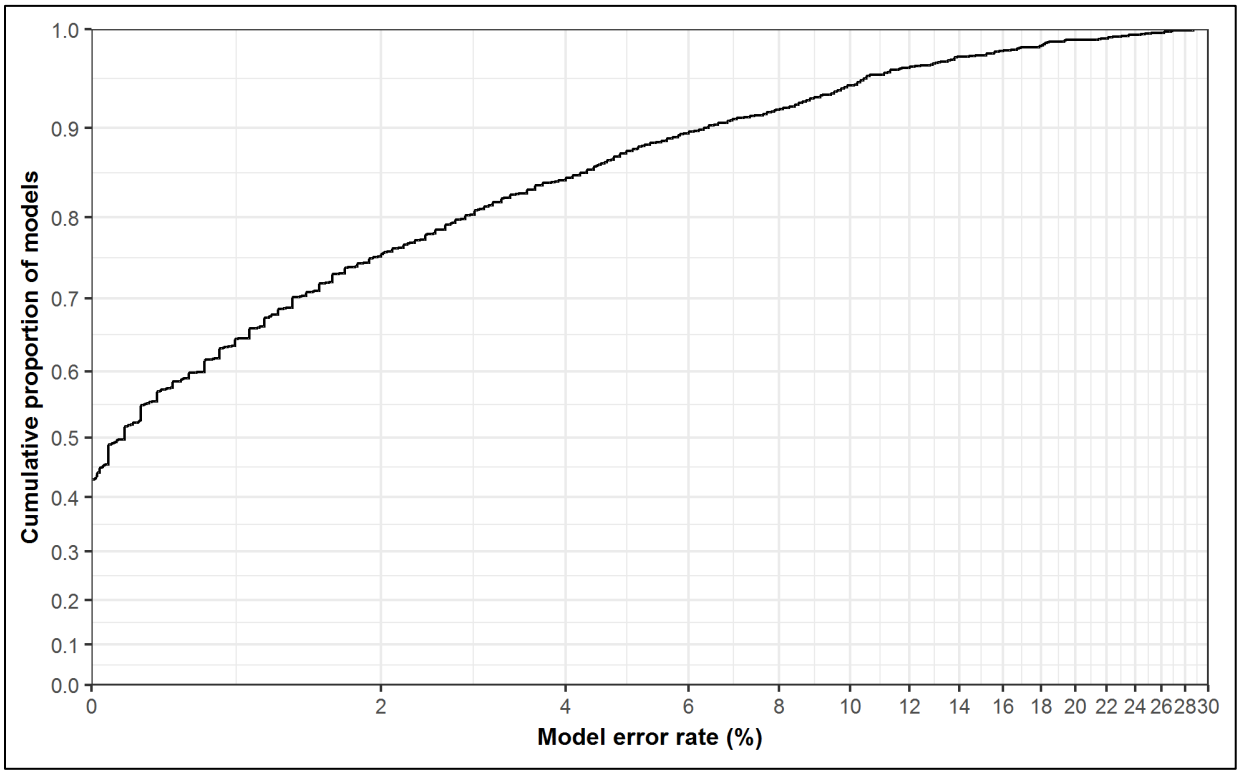


Figure 14. Error rate across all binary models(no-lidar run). Recall that due to heterogeneity within training samples (i.e., "inclusions"), some model error is to be expected.

Significant model error is highly concentrated in a fairly small number of binary one-versus-one models; these are map class distinctions that are more likely to map poorly. Although the accuracy assessment (**Section 3**) provides more definitive metrics of map accuracy, some map classes were poorly sampled in the accuracy assessment; for those, model error may be useful supplementary information. **Table 12** lists the 40 models with the highest error rates in the lidar run, used to provide map predictions over most of the park.

[60] When the classes to be predicted are not represented evenly in the training data, the more common class has a tendency to be modeled with greater accuracy than the other. This effect can be alleviated by downsampling the more common class (see Evans and Cushman 2009). The same technique was used earlier during the predictor selection phase.

Table 12. Binary models with highest cross-validated error rates(lidar run).

Map class 1 / Map class 2 (codes and abbreviated names)	Error rate (%)
H62–Alpine sparse herbaceous / R72–Colluvial barren	31.1
C05–W hemlock & sword fern / C06–W hemlock & salal	29.2
H50W–Lowland marsh & meadow / S40W–Low elevation shrub wetland	27.9
H52–Cow parsnip / H53–Showy sedge & Sitka valerian	24.3
C13–Mtn hemlock & Cascade azalea / C14–Silver fir & big huckleberry	24.2
H56–Subalpine summer-dry grass-forb / H61–Spreading phlox & common juniper	23.6
H54–Moist talus vegetation / R72–Colluvial barren	22.7
H60W–Black alpine sedge / S49–Alpine heather	22.1
B30–Successional gravel bar / R71–Alluvial barren	21.8
C04–Moist w hemlock & foamflower / C05–W hemlock & sword fern	21.8
C04–Moist w hemlock & foamflower / C10–Moist silver fir & foamflower	21.6
C12–Silver fir & Alaska blueberry / C14–Silver fir & big huckleberry	21.6
C05–W hemlock & sword fern / C11–Mesic silver fir & w hemlock	21.5
C13–Mtn hemlock & Cascade azalea / C20–Subalp fir & Sitka valerian	21.1
H53–Showy sedge & Sitka valerian / S47–Successional huckleberry	20.8
C06–W hemlock & salal / C08–Olympic dry Doug-fir	20.7
H60W–Black alpine sedge / S48–Subalpine heather	20.7
H61–Spreading phlox & common juniper / H62–Alpine sparse herbaceous	20.3
H52–Cow parsnip / S44–Thimbleberry, forbs & bracken fern	19.7
B30–Successional gravel bar / S40W–Low elevation shrub wetland	19.6
C03–Sitka spruce & wood-sorrel / C04–Moist w hemlock & foamflower	18.6
B30–Successional gravel bar / M90–Ruderal meadow	18.0
C10–Moist silver fir & foamflower / C11–Mesic silver fir & w hemlock	17.6
H55–Olympic aster scree meadow / H62–Alpine sparse herbaceous	17.3
S48–Subalpine heather / S49–Alpine heather	16.3
C05–W hemlock & sword fern / C08–Olympic dry Doug-fir	16.2
H51W–Subalpine herbaceous wetland / S41W–Subalpine willow wetland	16.2
R72–Colluvial barren / R73–Bedrock barren	15.9
H54–Moist talus vegetation / H55–Olympic aster scree meadow	15.8
C04–Moist w hemlock & foamflower / C11–Mesic silver fir & w hemlock	14.8
H54–Moist talus vegetation / H62–Alpine sparse herbaceous	14.8
H56–Subalpine summer-dry grass-forb / H60W–Black alpine sedge	14.7
C21–Mtn hemlock & heather / S48–Subalpine heather	14.6
H62–Alpine sparse herbaceous / R73–Bedrock barren	14.2
C12–Silver fir & Alaska blueberry / C13–Mtn hemlock & Cascade azalea	14.1
H54–Moist talus vegetation / R73–Bedrock barren	13.9
R71–Alluvial barren / W81–Fresh water	13.9

Table 12 (continued). Binary models with highest cross-validated error rates (lidar run).

Map class 1 / Map class 2 (codes and abbreviated names)	Error rate (%)
S44–Thimbleberry, forbs & bracken fern / S45–Vine maple	13.8
H51W–Subalpine herbaceous wetland / H53–Showy sedge & Sitka valerian	13.7
B31–Broadleaf riparian & swamp forest / B33–Upland bigleaf maple & conifer	13.5

2.7.6. Model prediction

The map class prediction at each 3-meter pixel was made by evaluating the results of each one-versus-one model and determining which class had the best overall performance. The "winner" of each model was determined using a simple threshold of 50% of the 507 trees. **Figure 15** shows the outcome of a single binary model in one small area. We accomplished this by creating a round-robin schedule of "contests" using the circle method of Reverend Kirkman (1847). Not all models needed to be evaluated at each pixel; after a map class had "lost" five contests, it was eliminated from contention and any subsequent models including it were skipped. The selection of the "winning" map class was made by comparing the total probability loss margin across all models, rather than by the number of contests lost, which removed the possibility of tie outcomes.

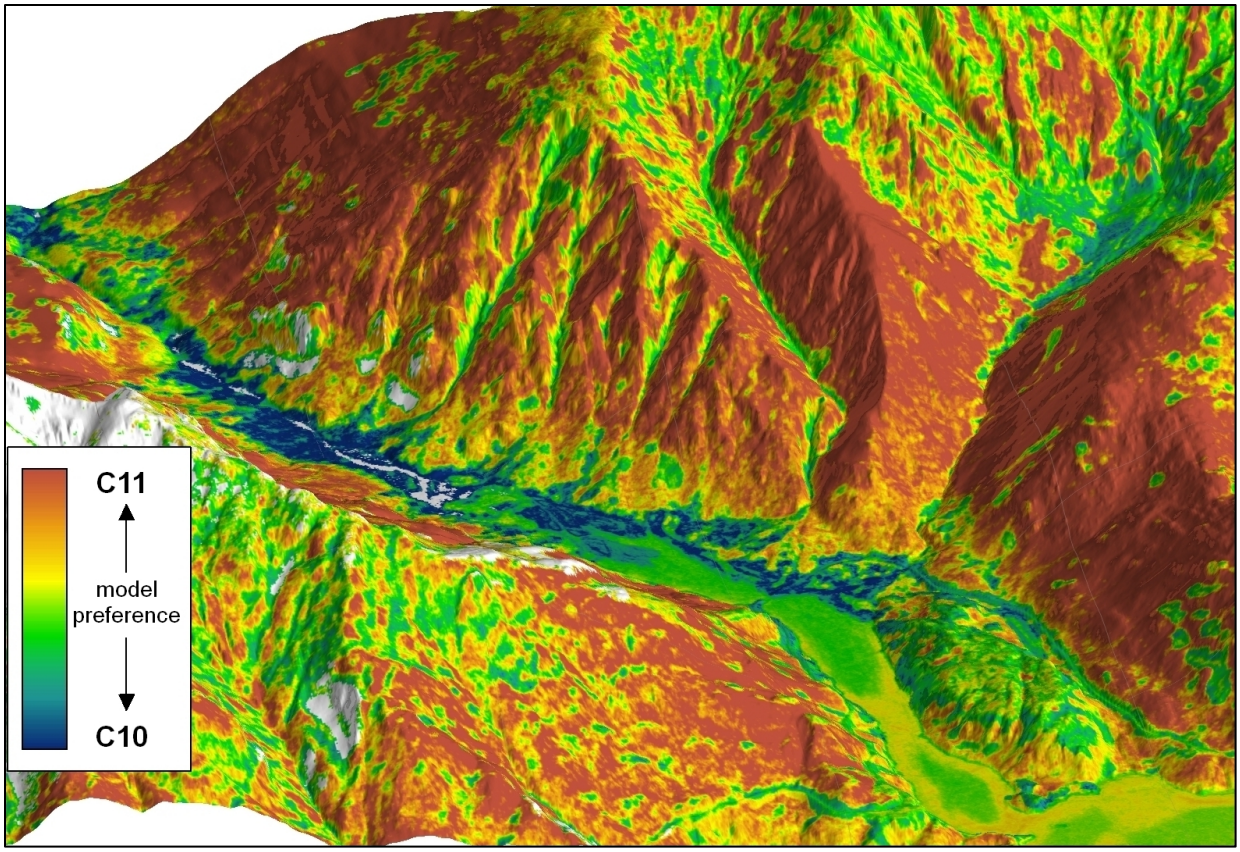


Figure 15. Binary model prediction example for the model C10–MOIST SILVER FIR, WESTERN HEMLOCK AND FOAMFLOWER FOREST versus C11–MESIC SILVER FIR AND WESTERN HEMLOCK FOREST. C10 was favored mostly on toe slopes and valley bottoms, while C11 was preferred on middle and higher slopes. White areas were not predicted because one or both map classes had already been eliminated from contention.

Processing made use of the R **randomForest** (Liaw and Wiener 2002), **raster** (Hijmans 2018), and **rgdal** (Bivand et al. 2014) packages, and was made more efficient by dividing the project area into tiles of approximately 2000 by 2000 pixels each. Each concurrent prediction process loaded the full set of predictor rasters for a single tile into memory and evaluated all needed models, tracking the number of losses and total loss margin by map class. The results for each binary model and the tracking information were copied to the network-attached storage device.[61] The total loss margin, seen in **Figure 16**, can be interpreted as a map of model uncertainty.

[61] Although multithreaded prediction (using all available CPU cores in a single process) is possible in R, we encountered reliability issues with this approach, and also found it was significantly more efficient to use multiple single-threaded processes.

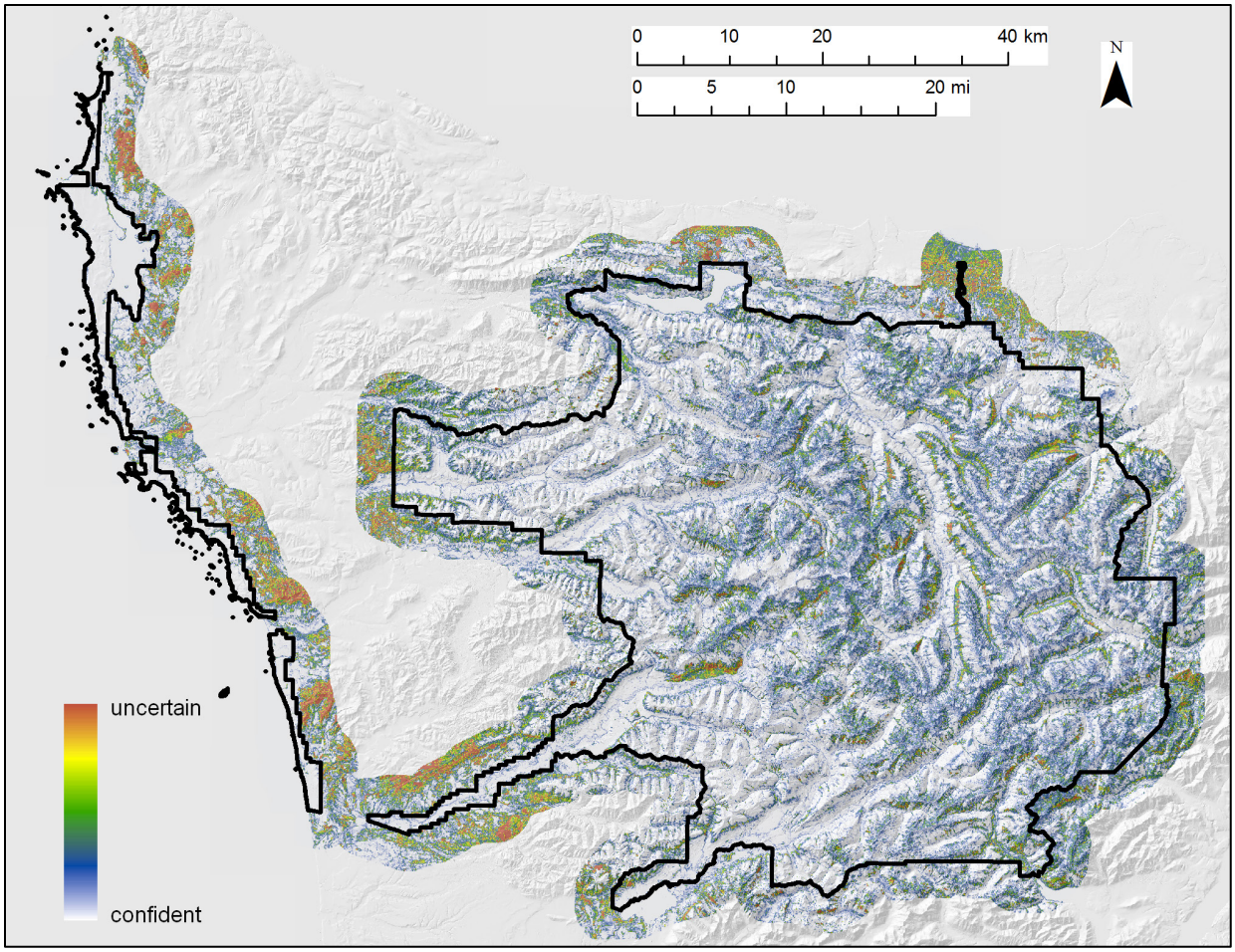


Figure 16. Prediction uncertainty **in the no-lidar run.** The predicted class won all contests in areas displayed as white. Colors from blue to red indicate that the best class lost at least one contest, by increasing amounts. Certainty is lowest where training data were inadequate, especially outside the park and in burned areas.

After predicting both the lidar and no-lidar runs, we created the final modeled map by using the lidar run result wherever it was available and the no-lidar result elsewhere (notably in the coastal strip).

2.8. Post-processing

The random forests pixel-based predictions were converted into a final map by means of a sequence of post-processing steps. We used various approaches, described below, to add additional map classes for land cover types within the project area but not represented in training data. A filtering process was then used to convert the pixel-based predictions into a polygon-based map. After that, we did a final phase of manual map editing to address observed problems in a few areas.

2.8.1. Additional map classes

We defined several additional map classes—M92–BURNED WITH UNCERTAIN VEGETATION, M93–TIMBERLAND WITH UNCERTAIN VEGETATION, M94–DEVELOPMENT and M95–ROADS IN PARK—to represent land cover types that were present in the project area but not in the training data.

Burned areas

The tasseled cap wetness index is particularly effective at estimating structural attributes in conifer-dominated forests (Cohen and Spies 1992). We created a mask of recently burned areas by subtracting the tasseled cap wetness calculated from the historic midsummer satellite imagery from that calculated from the current midsummer imagery (see **Sections 2.5.2** and **2.6.2**). We empirically determined a change threshold that was effective in flagging areas that had experienced severe burns between those dates, restricting the results to pixels within the digitized recent fires mask (**Section 2.1.2**). For areas identified as burned, if the model prediction was not a map class associated with early recovery from fire[62] (e.g., B33–UPLAND RED ALDER, BIGLEAF MAPLE AND CONIFER FOREST, S44–THIMBLEBERRY SHRUBLAND, TALL FORBS AND BRACKEN FERN, S47–SUCCESSIONAL HUCKLEBERRY SHRUBLAND), we recoded it to M92–BURNED WITH UNCERTAIN VEGETATION. The burned class included post-fire recovery areas as well as recent burns. We assumed that the model prediction would generally be an acceptable result for areas recovering from burns earlier than the mid-1980s. Because east-side post-fire communities are better represented in the classification and training data, burns in west-side forests are more likely to be mapped as M92–BURNED WITH UNCERTAIN VEGETATION.

Development

A variety of land cover types on both sides of the park boundary are actively maintained by human land-use practices (e.g., buildings for residential and commercial purposes, agriculture, roads). We designated two map classes to encompass these: M95–ROADS IN PARK (representing roads within the park boundaries only) and M94–DEVELOPMENT (representing everything else, including other development within the park).

We began by digitizing the roads and developed areas inside the park (using lidar data where available and 2015 NAIP otherwise), as well as major areas of development and agriculture outside the park boundary. Hand-mapping was done at 1:4,000 scale; road centerlines were buffered by either seven or 14 meters depending on their size. We created a managed areas mask by excluding the park and adjacent USFS wilderness from the project area, for use in later steps.

We added roads outside the park to M94–DEVELOPMENT by resampling the developed land cover classes in the 2016 National Land Cover Database (Yang et al. 2018) to our mapping resolution and removing areas that lay either outside the managed areas mask or within our digitized development and agriculture layer. Areas within the digitized development and agriculture layer were then added to M94–DEVELOPMENT unless they had modeled as a forest or woodland, a dense tall shrubland, or ocean, all of which can be reasonable map results in those places. Although ponds are a frequent feature in developed areas, we excluded them as there was a tendency to erroneously map fresh water in flat developed areas with cast shadows from adjacent trees. Digitized developed areas within the park were also included in M94–DEVELOPMENT. Digitized roads within the park were included as a separate map class, M95–ROADS IN PARK.

[62] We referred to Franklin and Dyrness (1973) in compiling this list.

Logging

We began by flagging disturbed forests, treating the impact of logging similarly to that of fires, by thresholding the historic change in tasseled cap wetness to detect areas that had experienced major canopy loss since the mid-1980s. We then applied a multi-stage majority filter and excluded areas that were smaller than a half-hectare, were within the park or adjacent wilderness areas, or had already been assigned to M94–DEVELOPMENT. Because flooding along major rivers was another significant cause of forest disturbance, areas that modeled as a typically riparian map class, were within five vertical feet of a major river and in a location with high hydrologic wetness (see **Section 2.6.4**) were also excluded.

The remaining disturbed areas were identified as potentially logged and were examined manually to remove those that did not appear to be within timber harvest boundaries. The rest were recoded to M93–TIMBERLAND WITH UNCERTAIN VEGETATION unless they had modeled as a forest, woodland or tall shrubland map class. The timberland class included early seral and planted forests, as well as recent regeneration harvests. We assumed that the model prediction would generally be an acceptable result for areas recovering from logging earlier than the mid-1980s.

2.8.2. Filtering

We converted the 3-meter pixel predictions to a polygon map via a sequence of filtering steps. Because lifeform can be predicted at very high accuracy but map classes are less easily distinguished, we began with a lifeform-specific majority filter that reassigned each pixel to the most common map class of the same lifeform among the neighboring pixels. No pixels were changed to a different lifeform than that to which they were predicted during this step. The analysis window ranged from 3-by-3 to 7-by-7 pixels depending on lifeform. We next addressed fine scale speckle by applying two successive 3-by-3 pixel majority filters across all map classes with no lifeform specificity.

We then moved to filtering based on patch size and shape, beginning by removing very small patches of fewer than nine contiguous 3-meter pixels, reassigning pixels in those patches to the nearest persisting patches. The shortest distance from each pixel to any neighboring patch was determined and the mean depth of each patch (d) was found by summarizing over its constituent pixels. Through experimentation, we defined an additional parameter g to describe patch shape:

$$g = d^{3/2}\, a^{-1/4}$$

where a is the patch area. While d describes the average width of a patch, g is a shape parameter describing the width of a patch compared to its overall size. We then empirically determined map class-specific thresholds for d and g; patches for which either parameter exceeded its threshold were kept, while the others were eliminated, assigning the constituent pixels to the nearest adjacent patch. This allowed us to filter map classes that often occur in long slender strips (e.g. C26–CONIFER KRUMMHOLZ AND TREED CLIFF) differently than those that typically occur over more extensive areas. We followed this with a final additional patch size filter, with a map class-specific size requirement ranging from nine to 49 pixels (81–441 m^2).

2.8.3. Map editing

We scanned the resulting polygon map for obvious errors, correcting them by hand by either recoding entire polygons or splitting them into pieces first. Issues that frequently required correction included occasional modeling of H50W–LOWLAND MARSH AND MEADOW in lake interiors near NAIP imagery join lines, confusion between M90–RUDERAL MEADOW and B30–SUCCESSIONAL GRAVEL BAR SHRUBLAND due to inadequate training data, and poor mapping along the edges of roads in the park. We then converted back to raster format and ran a final 3x3-pixel majority filter to eliminate any stray missing pixels from the map before converting back to polygons for the final time. The final map is available in Nielsen et al. (2021d).

3. Accuracy Assessment

3.1. Background

A map accuracy assessment (AA) determines the degree to which a map correctly represents on-the-ground conditions (see Lea and Curtis 2010, Congalton and Green 1999). A *confusion matrix* or *contingency table* tabulates the misassignments between each possible pair of map classes. The information from this matrix is used to draw conclusions about the quality of mapping for each map class; the results allow an evaluation of potential map applications and applicable caveats. *User's accuracy* (UA) and *producer's accuracy* (PA) describe two relevant aspects of map accuracy.

UA is a reliability measure to estimate the probability that the map is correct where a particular class is mapped. It is inversely related to the false-positive or *commission error* (CE) rate (the probability of mapping the class where it is not present). Low UA may indicate that a class is *over-mapped* (mapped more often than it actually occurs). It also can be evidence that classes that are particularly confused with it are themselves *under-mapped* (mapped less often than they actually occur).

PA is a mappability measure to estimate the probability that the map is correct where a particular class is found on the ground. It is inversely related to the false-negative or *omission error* (OE) rate (the probability of omitting the class where it is present). Low PA may indicate that a class is under-mapped. Because PA is relative to the true land cover, rather than the mapped land cover, its calculation is dependent on an estimate of the true quantity of the class present in the study area. Thus, two distinct estimates of PA can be made. The first, relative to the number of plots found in the field, is calculated from a confusion matrix drawn directly from the sampled plots, the *sample contingency table* (SCT). The second, a more meaningful quantity, is scaled to an estimate of the true area occupied by each class, the *population contingency table* (PCT).

We followed the procedures and formulas provided by Lea and Curtis (2010) for sample design, sample protocol, and analysis, to the extent possible.[63] NPS standards specify an 80% accuracy goal

[63] The logistics involved in mounting the AA field campaign with an experienced field crew—in addition to other project management considerations—required the AA fieldwork to occur several years earlier than would have been ideal. At the time of the fieldwork, we had only a very early version of the map classification, and were still relying on the provisional NVC associations and the accompanying keys (Crawford et al. 2009), which were later discovered to have a number of issues (necessitating the creation of the mapping associations). The early draft map we relied on for our sample design in some areas bears little resemblance to the final map. Because we recognized the challenges that lay ahead in making the AA data compatible with an as-yet uncompleted classification and map, we had no choice but to violate some important guidance from Lea and Curtis (2010), collecting substantial field data during the fieldwork, and using them later in the office to arrive at a best call using the final classification (though we did lean heavily on the field crew's call whenever possible). Cognizant of the concern that the expertise possessed by team members would not be available to later users, we based our assessment solely on field materials that will be available—namely, the map class indicator species lists available in Nielsen et al. (2021c). Furthermore, all map production steps undertaken subsequent to the fieldwork—the creation of mapping associations, their crosswalk to map classes, changes in the modeling scheme, and (most importantly) post-processing and map

for each individual map class hosting native vegetation communities. In addition to assessing the class-specific UA and PA against this standard, we used UA and PA in combination to produce an estimate of the true area occupied by each class in the park. These estimates were in turn used to adjust the overall map accuracy by area-weighting the per-class accuracies; this step was necessary to compensate for the stratified random sampling design that guided sample selection.

3.2. Sample design

We used a random sampling approach, stratified by mapped class, to select sample locations. We targeted all natural vegetated classes mapped within the park except for M92–BURNED WITH UNCERTAIN VEGETATION, which we excluded due to its rarity and difficulty of access. We also targeted two nominally abiotic classes, R71–ALLUVIAL BARREN AND DEBRIS-COVERED ICE and R72–COLLUVIAL BARREN, because they often host a small amount of vegetation and have potential for confusion with several vegetated map classes, such as B30–SUCCESSIONAL GRAVEL BAR SHRUBLAND and H62–ALPINE SPARSE HERBACEOUS VEGETATION. Accurate mapping of these abiotic classes is necessary in order to accurately map similar vegetated classes. While R73–BEDROCK BARREN also has the potential for confusion with vegetated map classes, it is difficult and often unsafe to access, and we deemed it not a good use of field time to attempt it. All map classes were mapped on greater than 50 hectares, corresponding to a sampling goal of 30 plots each (Lea and Curtis 2010).

3.2.1. Inference area

The AA inference area was defined by buffering park roads and trails by 200 meters.[64] This narrow buffer was necessitated by spatial inaccuracies in the park trails data and the recognition that the true distance from the trail might be as large as the summed buffer distance and the local trail spatial inaccuracy. We chose not to use a slope-based cost distance model or further limit the inference area based on slope requirements because of concerns about the spatial accuracy of both the trails layer and the topographic data (derived from the USGS DEM since AA was done before the lidar data had been collected). The resulting region (**Figure 17**) spanned 31,446 hectares, 8.5% of the park's total area.[65] Because crews were not able to visit the full area in which sample locations were generated, we will refer to the targeted sampling region as the *attempted inference area* (AIA).

---

editing—were done without reference to the AA field data, which were not fully processed until the final year of the project.

[64] Restricting the sampling area to road and trail corridors was necessary for safety and efficiency. Some segments of the trail network, including a portion of the Dosewallips River, Boulder Creek and Happy Lake Ridge, Hayden Pass, Skyline Ridge, and others, were not included in the AA effort. The single season available for sampling required that backpacking trips be planned for maximum efficiency, and these trail segments would have required too much effort to reach or did not fit well with other established sampling objectives.

[65] Although this falls well short of the 30% standard given in Lea and Curtis (2010), it was impossible to achieve sampling productivity goals with a larger inference area (see **Table 13** for a summary of the per-class inference area proportion). Given limited resources, we chose to prioritize the number of samples collected over expanding the inference area.

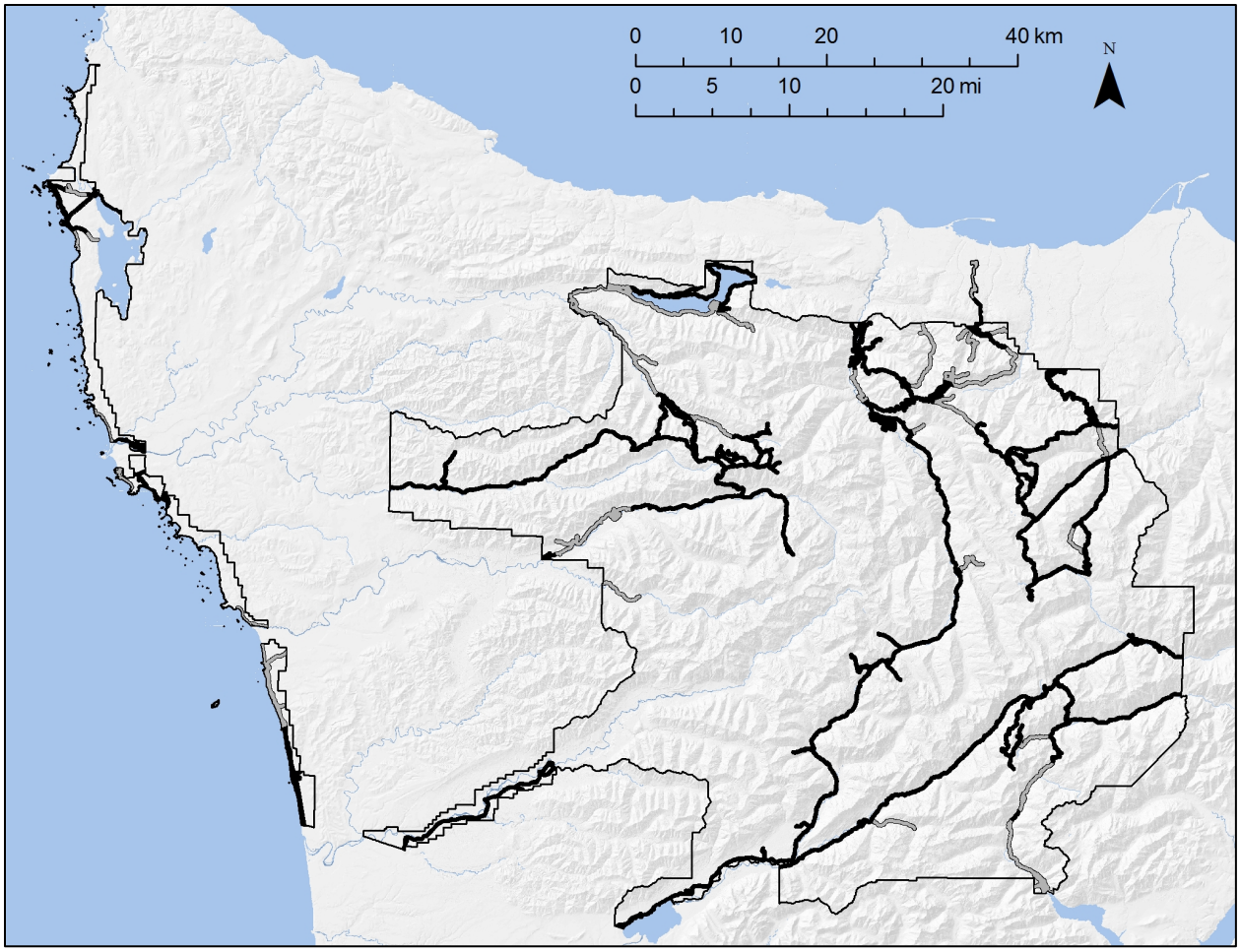


Figure 17. Accuracy assessment inference area Field samples were acquired within the reached inference area, shown in black. The attempted inference area also includes gray-shaded areas which were not accessed by field crews.

3.2.2. Sample selection

We attempted to minimize map class membership ambiguity in the samples by selecting them from locations that were at least 40 meters from a pixel mapped to any other class (the radius corresponding to the NPS standard half-hectare MMU). Increasing the buffer further (e.g., to allow for field and map positional error, see Lea and Curtis 2010) would have not permitted collecting adequate samples of many of the map classes. To make obtaining sufficient samples for each stratum more likely, considering the limitations on available crew time in the field and the likelihood of inaccessible samples, we generated twice as many target points for each mapped class than the stratum-specific sampling goal. Because crew size limitations prevented independent observations for overlapping plots, we specified a minimum separation of 60 meters between targets, irrespective of mapped class. There remained the possibility that plots might overlap, but any larger separation would, again, have made collecting adequate samples impossible for some classes.

We assigned a random number to each of the 60 points generated for each class and coded the targets with the 30 lowest numbers within each class as required plots, with the other 30 being optional.

Crews were to plan their field logistics around collecting the required plots for each map class, with additional accessible optional plots encountered being sampled if time permitted. This allowed us to obtain a spatially representative collection of plots from each class—spreading samples across different field trips—without lowering the chances of obtaining an adequate number of samples for any. If required plots were found to be inaccessible, additional required plots were designated from the remaining optional plots to maintain the goal of 30 plots per class.

3.3. Field data

3.3.1. Field logistics

INR and NPS field crews collected field data from July 15 to September 27, 2013. A total of 11 seven-day tours were accomplished by crews of two to four members (one or two groups, of one or two members each). As crews moved along roads and trails, they sampled every required plot that the day's logistics and safety factors would allow and spent any additional time available sampling optional plots along the route. After each tour, a day was spent in the office entering a brief summary of each plot along the route: whether it was attempted, whether it was reached, and what the field map class call had been. This summary was used by INR office staff to assess the updated prognosis for achieving the sampling goal for each map class, based on the opportunities remaining in the target pool. By the time field crews returned for their next tour, particular plots had been identified as high priority (based on being mapped as map classes that were most in danger of falling short of sampling goals) and crew leaders planned their itineraries accordingly.[66] Missions planned for subsequent weeks were adjusted on an as-needed basis as some trips contained a greater proportion of high priority map classes.

3.3.2. Field protocol

Crews navigated to each target location and assessed the surroundings. If transitions to alternate map classes were nearby, they moved the plot center to a more homogeneous point in the same vegetation type and updated the location using a GPS. If plot centers were not safely accessible and high confident assessments of plot location and vegetation call could be made, crews were permitted to make their observations from a distance. In this case, they noted the distance and bearing at which the plot lay from their observation point. Plots that could not be reached or assessed were discarded.

We used a half-hectare assessment area for all vegetation types, corresponding to a 40-meter assessment radius around plot center. Crews identified the plant associations (if their level of experience permitted) or map classes (otherwise) covering at least 20% of the plot area.[67] Multiple calls occurred frequently in small patch vegetation types dominated by herbaceous plants and dwarf shrubs. Although spatial heterogeneity was less of an issue for most forest plots, ambiguity in association and even map class were occasionally encountered, usually due to intermediate

[66] Field crews were not knowledgeable of the class that each plot had been mapped to, to avoid influencing their judgment.

[67] The plant association descriptions and keys (Crawford et al. 2009) and descriptions of the nascent map classes, which later evolved into Nielsen et al. (2021c), were provided for this purpose.

combinations of understory species or borderline composition situations (e.g., *Abies amabilis* at or near 10% canopy cover). Up to four associations or map classes with a plausible fit to the vegetation present were recorded. The crew's confidence in each call was recorded as high, medium or low. Although unvegetated map classes were not targeted, they were permitted as field calls.

A variety of data were collected to assist with subsequent quality control. Percent cover estimation for a fixed set of important indicator species and unvegetated ground components, including all tree species and a variety of shrubs, forbs and graminoids, allowed for subsequent checks on the vegetation calls. Tree species were broken into canopy, sub-canopy, and regeneration layer components. The identities of the five most prevalent species in both the shrub and herbaceous layers were also recorded. Crews noted the topographic setting and sketched a diagram showing nearby reference points to provide additional confidence in assessing plot location. Finally, they provided a brief description of the plot and pros and cons for each of their association or map class calls. **Figure 18** contains a completed AA field form.

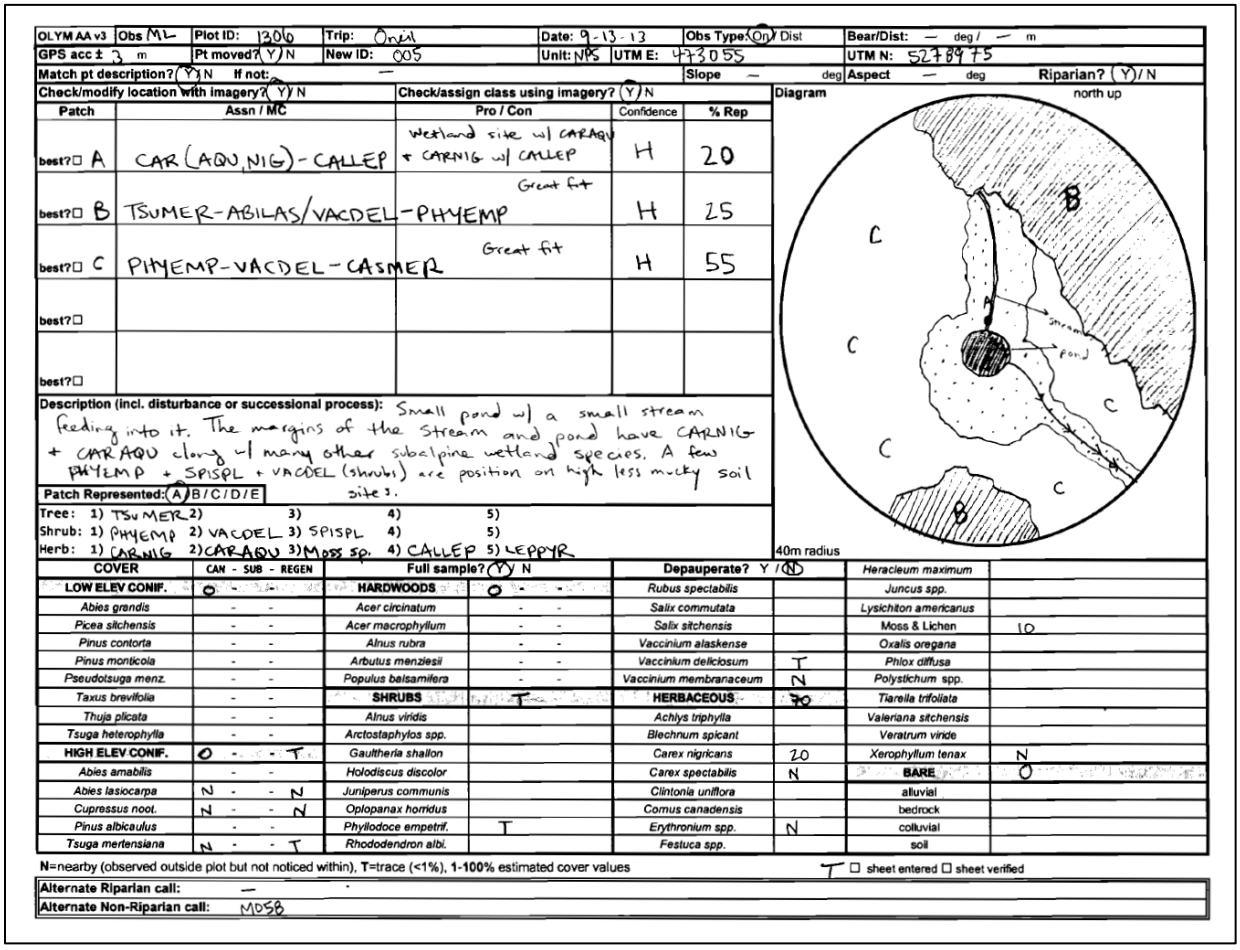


Figure 18. Completed accuracy assessment data sheet collected in 2013.

3.3.3. Quality control

Accuracy assessment plot data went through a quality control process similar to that of map training data. The NPS phase of the spatial accuracy QC started after the field tour. At the conclusion of each tour, locations representing moved points were uploaded and reviewed for accuracy by a team member who had visited the plot. Subsequent INR QC verified and if necessary adjusted the location by comparing NAIP imagery and the field diagram.

Using the species cover data, plot description and diagram, and with reference to imagery and map class indicator species (see **Section 2.5.3**), each plot was checked for the correctness of the map class calls. Every effort was made to label each plot with a single best map class call. However, there were two situations in which we chose to allow some flexibility. Sometimes, we were unable to make a single best call because the floristics field data were perfectly intermediate between two classes. Both calls were treated as legitimate possible answers at the 42 plots where this occurred. In another 45 plots, the identified homogeneous patches were so small that we anticipated the possibility of a label mismatch when in fact the vegetation was correctly predicted. This might result from spatial offset between the GPS location, the layers by which the analyst assessed the plot, and the model-based prediction, as well as from filtering the predictions to MMU scale. If the boundary with another map class was within ten meters or less of the assessed point, we entered that map class as an additional correct answer.

Throughout this process, the analyst did not have access to the final map polygons or labels. If a plot could not be confidently located or its correct map class could not be narrowed down to at most two possibilities, it was discarded. Forty-five plots were rejected on these grounds, most of them because of uncertain location, extreme heterogeneity, or ambiguity due to the impact of disturbance.

3.3.4. Field plot totals and reached inference area

A total of 926 field plot samples were collected for the accuracy assessment; 881 of them passed the QC process. We determined the *reached inference area* (RIA) by removing from the AIA any route portions at least 1 km long and lacking sampled plots. The RIA totaled 23,945 hectares, equivalent to 76.1% of the attempted inference area or 6.5% of the full park. **Table 13** gives the overall effectiveness at reaching the accessible portions of each field-targeted map class and the class-specific inference area fraction. We fell well short of the standard of Lea and Curtis (2010), who specify that "a minimum of at least the most accessible 30th percentile of abundant classes should be included in even the most difficult of access situations." That was achieved for only two classes, one of which was the rarest class in the park, M90–RUDERAL MEADOW.

Table 13. Map class-specific accuracy assessment inference areas For each class, the mapped area in (a) the park, (b) the attempted inference area (AIA) and (c) the reached inference area (RIA), followed by (d) the fraction reached of the area mapped in the AIA (a measure of field effectiveness), and (e) the fraction of the total mapped area represented in the RIA (a measure of the representativeness of the inference area).

Class code and abbreviated name	Mapped in park (ha)	Mapped in AIA (ha)	Mapped in RIA (ha)	% of AIA reached	% Rep in RIA
C11–Mesic silver fir & w hemlock	59,071	1,199	945	78.9	1.6
C12–Silver fir & Alaska blueberry	24,553	664	592	89.2	2.4
C26–Conifer krummholz & treed cliff	5,749	188	160	85.4	2.8
H54–Moist talus vegetation	1,185	44	34	78.8	2.9
R72–Colluvial barren	9,528	345	307	88.9	3.2
H62–Alpine sparse herbaceous	3,116	111	106	96.0	3.4
H55–Olympic aster scree meadow	3,996	172	159	92.3	4.0
C13–Mtn hemlock & Cascade azalea	23,398	1,096	956	87.2	4.1
H58–Bedrock balds and forest openings	518	28	21	75.6	4.1
C21–Mtn hemlock & heather	18,783	796	789	99.1	4.2
C14–Silver fir & big huckleberry	7,285	435	340	78.0	4.7
S43–Sitka alder	4,964	309	233	75.5	4.7
C06–W hemlock & salal	30,128	2,250	1,454	64.6	4.8
C10–Moist silver fir & foamflower	33,633	2,075	1,666	80.3	5.0
C15–Lodgepole pine & Doug-fir	1,109	94	58	61.7	5.2
S47–Successional huckleberry	1,659	104	103	99.5	6.2
S44–Thimbleberry, forbs & bracken fern	1,059	82	68	82.2	6.4
C20–Subalp fir & Sitka valerian	17,498	1,369	1,138	83.2	6.5
C00–Sitka spruce, salal & sword fern	5,360	672	415	61.7	7.7
H60W–Black alpine sedge	295	24	23	96.0	7.9
C01–Redcedar, salal & evergreen huckleberry	6,500	895	529	59.1	8.1
S49–Alpine heather	1,155	107	99	93.3	8.6
H53–Showy sedge & Sitka valerian	1,994	186	174	94.0	8.7
C04–Moist w hemlock & foamflower	19,314	2,635	1,866	70.8	9.7
C24–Olympic subalp fir & lodgepole pine	4,737	563	468	83.2	9.9
H51W–Subalpine herbaceous wetland	1,011	105	103	97.7	10.2
S45–Vine maple	833	112	87	77.7	10.4
S48–Subalpine heather	3,015	321	316	98.3	10.5
C03–Sitka spruce & wood-sorrel	20,962	2,878	2,342	81.4	11.2
B32–Coastal red alder	790	229	90	39.3	11.4
C05–W hemlock & sword fern	12,006	2,630	1,649	62.7	13.7
C02–Redcedar & sphagnum bog	210	58	31	52.9	14.5
S41W–Subalpine willow wetland	720	125	115	91.8	16.0

Table 13 (continued). Map class-specific accuracy assessment inference areas. For each class, the mapped area in (a) the park, (b) the attempted inference area (AIA) and (c) the reached inference area (RIA), followed by (d) the fraction reached of the area mapped in the AIA (a measure of field effectiveness), and (e) the fraction of the total mapped area represented in the RIA (a measure of the representativeness of the inference area).

Class code and abbreviated name	Mapped in park (ha)	Mapped in AIA (ha)	Mapped in RIA (ha)	% of AIA reached	% Rep in RIA
S42–Sitka willow riparian	423	85	69	81.2	16.2
H52–Cow parsnip	462	86	77	89.0	16.6
C08–Olympic dry Doug-fir	5,676	1,369	988	72.2	17.4
R71–Alluvial barren	2,626	596	456	76.5	17.4
S40W–Low elevation shrub wetland	118	24	23	94.2	19.5
H50W–Lowland marsh & meadow	350	81	71	87.1	20.2
H61–Spreading phlox & common juniper	1,690	438	386	88.0	22.8
B33–Upland bigleaf maple & conifer	2,328	767	569	74.2	24.4
B30–Successional gravel bar	920	277	231	83.5	25.1
H56–Subalpine summer-dry grass-forb	495	137	134	97.8	27.0
B31–Broadleaf riparian & swamp forest	6,050	2,187	1,857	84.9	30.7
M90–Ruderal meadow	62	26	26	100.0	40.9

3.4. Photo-interpretation

An additional 110 plots were randomly generated across the ranges of five mapped abiotic classes that were not targeted for field sampling (R70–COASTAL BARREN, R73–BEDROCK BARREN, W80–SALT WATER, W81–FRESH WATER and W82–EXPOSED SNOW AND ICE). These points were photo-interpreted to map class. Although not included in the AA contingency tables, these classes occupy substantial portions of the park and we felt it was important to have an estimate of their accuracy.

3.5. Sampling outcomes

A total of 991 sample plots passed the QC process or were photo-interpreted (**Figure 19**). The final coordinates for each plot were used to extract the predicted map class label from the final map. **Table 14** gives, for each targeted map class, the mapped area (in hectares), the number of plots mapped as and identified as the class, and the fraction of the sampling goal that was achieved. Small numbers of mapped plots result in uncertain estimates of user's accuracy, whereas small numbers of identified plots result in uncertain estimates of producer's accuracy (and uncertain map area corrections for any classes confused with it).

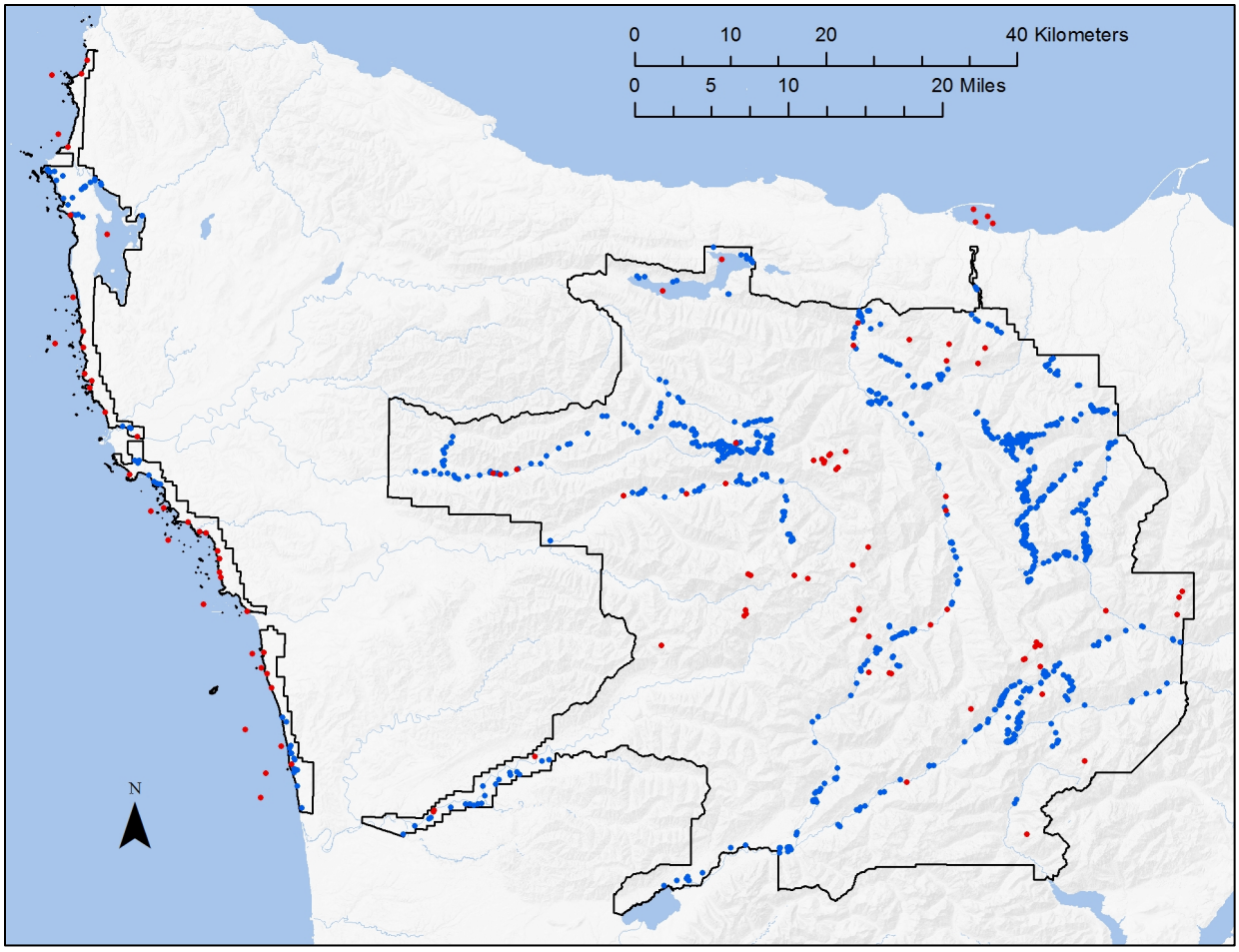


Figure 19. Accuracy assessment plot locations 881 quality-controlled field-collected AA plots (blue dots) and 110 photo-interpreted points for unvegetated map classes (red dots) are shown.

Table 14. Accuracy assessment plot totals For each class, the mapped area in the park, the number of AA plots mapped as and identified as the class, and the fraction of the sampling goal achieved (minimum based on mapped and identified plots). The sampling goal was 30 plots per class. Poorly sampled classes are listed first.

Class code and abbreviated name	Mapped area in park (ha)	# of plots mapped as	# of plots identified as	% of goal achieved
C15–Lodgepole pine & Doug-fir	1,109	3	0	0
H54–Moist talus vegetation	1,185	4	2	7
H58–Bedrock balds and forest openings	518	2	2	7
S40W–Low elevation shrub wetland	118	2	3	7
H60W–Black alpine sedge	295	5	5	17
H62–Alpine sparse herbaceous	3,116	5	6	17
S42–Sitka willow riparian	423	6	5	17
B32–Coastal red alder	790	6	7	20
C02–Redcedar & sphagnum bog	210	6	6	20

Table 14 (continued). Accuracy assessment plot totals. For each class, the mapped area in the park, the number of AA plots mapped as and identified as the class, and the fraction of the sampling goal achieved (minimum based on mapped and identified plots). The sampling goal was 30 plots per class. Poorly sampled classes are listed first.

Class code and abbreviated name	Mapped area in park (ha)	# of plots mapped as	# of plots identified as	% of goal achieved
C13–Mtn hemlock & Cascade azalea	23,398	12	8	27
C14–Silver fir & big huckleberry	7,285	16	9	30
H50W–Lowland marsh & meadow	350	12	9	30
C00–Sitka spruce, salal & sword fern	5,360	10	13	33
M90–Ruderal meadow	62	10	13	33
S45–Vine maple	833	13	12	40
S49–Alpine heather	1,155	13	12	40
H51W–Subalpine herbaceous wetland	1,011	14	13	43
R72–Colluvial barren	9,528	13	14	43
C11–Mesic silver fir & w hemlock	59,071	15	14	47
H55–Olympic aster scree meadow	3,996	15	14	47
B30–Successional gravel bar	920	15	15	50
C26–Conifer krummholz & treed cliff	5,749	16	15	50
S44–Thimbleberry, forbs & bracken fern	1,059	15	20	50
C12–Silver fir & Alaska blueberry	24,553	16	25	53
C01–Redcedar, salal & evergreen huckleberry	6,500	18	19	60
C08–Olympic dry Doug-fir	5,676	18	20	60
R71–Alluvial barren	2,626	19	18	60
B33–Upland bigleaf maple & conifer	2,328	19	20	63
H52–Cow parsnip	462	24	20	67
C21–Mtn hemlock & heather	18,783	22	31	73
S41W–Subalpine willow wetland	720	27	23	77
C04–Moist w hemlock & foamflower	19,314	24	25	80
C06–W hemlock & salal	30,128	26	25	83
C24–Olympic subalp fir & lodgepole pine	4,737	29	25	83
S43–Sitka alder	4,964	25	26	83
C10–Moist silver fir & foamflower	33,633	33	30	100
B31–Broadleaf riparian & swamp forest	6,050	32	31	103
S47–Successional huckleberry	1,659	33	31	103
C20–Subalp fir & Sitka valerian	17,498	35	33	110
S48–Subalpine heather	3,015	33	41	110
C05–W hemlock & sword fern	12,006	34	34	113
H53–Showy sedge & Sitka valerian	1,994	35	34	113

Table 14 (continued). Accuracy assessment plot totals. For each class, the mapped area in the park, the number of AA plots mapped as and identified as the class, and the fraction of the sampling goal achieved (minimum based on mapped and identified plots). The sampling goal was 30 plots per class. Poorly sampled classes are listed first.

Class code and abbreviated name	Mapped area in park (ha)	# of plots mapped as	# of plots identified as	% of goal achieved
H56–Subalpine summer-dry grass-forb	495	36	34	113
H61–Spreading phlox & common juniper	1,690	36	39	120
C03–Sitka spruce & wood-sorrel	20,962	57	57	190

The sampling goal of 30 samples per mapped class, needed for a confident assessment of UA, was achieved for only ten of the 45 classes included in the AA. Four classes (H58–BEDROCK BALDS AND SPARSELY VEGETATED FOREST OPENINGS, S40W–LOW ELEVATION SHRUB-DOMINATED WETLAND, C15–LODGEPOLE PINE AND DOUGLAS-FIR WOODLAND and H54–MOIST TALUS VEGETATION) had fewer than five mapped occurrences sampled. The success rate for identified plots, needed for a confident assessment of PA, was similar. Only 11 of the classes were identified at least 30 times. The same four classes that were poorly sampled from the UA perspective were also poorly sampled for PA. One class, C15–LODGEPOLE PINE AND DOUGLAS-FIR WOODLAND, was not identified at all.

On the other hand, some classes were oversampled, although for only one was this drastic (C03–SITKA SPRUCE, WESTERN HEMLOCK AND WOOD-SORREL FOREST). Although the plan had been to remove plots corresponding to types that had been adequately sampled from the crew's target lists, there were difficulties in implementing the plan given limited office time between field tours. Crews themselves were not able to steer away from oversampled types because they were unaware of the map class label attached to each point.

3.6. Analysis

A total of 45 classes were included in the AA analysis, including all classes hosting natural vegetation communities. The nominally abiotic classes R71–ALLUVIAL BARREN AND DEBRIS-COVERED ICE and R72–COLLUVIAL BARREN were also included; field sampling permitted a confident assessment of their accuracy and they often provide important habitat both for unmappably sparse plant communities and for wildlife populations. The five classes for which most AA samples were photo-interpreted[68] were excluded from further analysis. We noted that the 128 samples mapped as these classes indicated a user's accuracy of 98.4%, but felt that because they were not field-sampled and generally do not host natural vegetation communities, their inclusion would artificially inflate the overall accuracy. The four classes that were mapped via PI or deductive modeling[69] were also

[68] R70–COASTAL BARREN, R73–BEDROCK BARREN, W80–SALT WATER, W81–FRESH WATER and W82–EXPOSED SNOW AND ICE.

[69] M92–BURNED WITH UNCERTAIN VEGETATION, M93–TIMBERLAND WITH UNCERTAIN VEGETATION, M94–DEVELOPMENT and M95–ROADS IN PARK.

excluded from further analysis. Four field-sampled plots fell into one of these classes; all were correctly mapped. A total of 859 plots, all field-collected, remained in the analysis.

The predicted map class was extracted from the 3-meter pixel at each plot center point and compared to the field-sourced map class calls. For plots with two best calls, we allowed either as a correct answer; 11 of these 42 plots were counted as correct because the map matched the second of the two best calls. For plots with a secondary patch call within ten meters of the assessed center point, we also allowed that as a correct answer.[70] Of these 45 plots, seven were called correct based on matching the secondary patch call. If the plot was counted as incorrect, it was labeled as the first best map class call from the primary patch.

The sample contingency table was created by indexing the observed map class against the predicted map class for each plot and summing across all plots. User's accuracy was calculated for each map class by dividing the number of correct samples by the total number of samples mapped as that class. Overall sample-based map accuracy was calculated by dividing the total number of correct calls by the total number of samples; however, this measure is misleading as it is biased by the use of the stratified random sampling design, which does not sample map classes in proportion to their prevalence in the project area (sample-based PA are similarly biased). The resulting SCT is shown in INR (2021a). The overall accuracy based on the raw samples is 86.0%.

To address the bias introduced by the stratified random sample design, a population contingency table (PCT) was created by reweighting the proportions represented by the cells in each row of the SCT by the fraction of the reached inference area mapped to that class. Each cell of the PCT, rather than containing raw sample counts, represents the estimated proportion of the RIA that is mapped as the class defined by the cell's row and identified as the class defined by the cell's column. We recalculated PA and overall accuracy based on the PCT; the revised measures represent the best estimates of the results that would have been obtained had the AA sample design been based on a simple random sample. The resulting PCT is shown in INR (2021a). The overall accuracy, after correcting for map class prevalence in the inference area, is 84.1%. Note that the mapped areas in the table sum to 22,324 hectares rather than the 23,945 hectares contained in the RIA. The nine classes excluded from the analysis were mapped on the remaining portion of the RIA.

The kappa coefficient, which provides an accuracy measure that accounts for the probability of map class agreement by chance alone, was calculated. Kappa is a proportion ranging from 0–100%, where a value of zero indicates a map that is no more accurate than would be expected by chance alone. 90% confidence intervals were calculated for all accuracy estimates. Finally, a corrected area estimate was created for each map class by multiplying the sum of the proportions in each column by the total mapped area of all the map classes in the PCT.

The class-specific user's accuracies are summarized in **Table 15**. The accuracy estimate met the 80% standard for 34 of the 45 map classes. However, for only 16 was the 90% confidence interval entirely

[70] This was intended to partially address nonthematic errors resulting from spatial misregistration (Foody 2002).

above the 80% mark. Most of the 11 that failed to meet the standard were poorly sampled. Only two classes, C14–SILVER FIR, BIG HUCKLEBERRY AND BEARGRASS FOREST and C15–LODGEPOLE PINE AND DOUGLAS-FIR WOODLAND, were conclusively demonstrated to fail to meet the standard via the 90% confidence interval. For the latter class, only three samples were collected, all outside its main mapped range.

Table 15. Map class-specific user's accuracy (UA)for each assessed map class, with poorly mapped classes first. Asterisks indicate true values that are at least 90% confident to lie either fully above or below the 80% accuracy target.

Class code and full name	# of plots mapped as	UA estimate	UA 90% conf int
C15–Lodgepole pine and Douglas-fir woodland	3	0%*	0–17%
C14–Silver fir, big huckleberry and beargrass forest	16	50%*	26–74%
H54–Moist talus vegetation	4	50%	0–100%
C13–Mountain hemlock, silver fir and Cascade azalea forest	12	58%	31–86%
S42–Sitka willow riparian shrubland	6	67%	27–100%
C11–Mesic silver fir and western hemlock forest	15	67%	43–90%
H62–Alpine sparse herbaceous vegetation	5	80%	41–100%
H60W–Black alpine sedge wetland	5	80%	41–100%
H51W–Subalpine herbaceous wetland	14	71%	48–95%
C06–Western hemlock, Douglas-fir and salal forest	26	73%	57–89%
C05–Western hemlock, Douglas-fir and sword fern forest	34	74%	60–87%
H50W–Lowland marsh and meadow	12	75%	50–100%
H52–Cow parsnip meadow	24	79%	63–95%
C20–Subalpine fir and Sitka valerian forest and woodland	35	80%	67–93%
H55–Olympic aster scree meadow	15	80%	60–100%
S49–Alpine heather shrubland	13	85%	64–100%
H53–Showy sedge and Sitka valerian meadow	35	83%	71–95%
C04–Moist western hemlock, Douglas-fir and foamflower forest	24	83%	69–98%
C10–Moist silver fir, western hemlock and foamflower forest	33	85%	73–97%
S41W–Subalpine willow wetland	27	85%	72–98%
C24–Olympic subalpine fir and lodgepole pine woodland	29	86%	74–98%
C08–Olympic dry Douglas-fir forest	18	89%	74–100%
H58–Bedrock balds and sparsely vegetated forest openings	2	100%	75–100%
S40W–Low elevation shrub-dominated wetland	2	100%	75–100%
S47–Successional huckleberry shrubland	33	88%	77–99%
S45–Vine maple shrubland	13	92%	76–100%
H56–Subalpine summer-dry grass-forb meadow	36	89%	79–99%
C03–Sitka spruce, western hemlock and wood-sorrel forest	57	89%*	82–97%
S44–Thimbleberry shrubland, tall forbs and bracken fern	15	93%	79–100%

Table 15 (continued). Map class-specific user's accuracy (UA) for each assessed map class, with poorly mapped classes first. Asterisks indicate true values that are at least 90% confident to lie either fully above or below the 80% accuracy target.

Class code and full name	# of plots mapped as	UA estimate	UA 90% conf int
B30–Successional gravel bar shrubland	15	93%	79–100%
C26–Conifer krummholz and treed cliff	16	94%*	81–100%
C01–Redcedar, western hemlock, salal and evergreen huckleberry forest	18	94%*	83–100%
B33–Upland red alder, bigleaf maple and conifer forest	19	95%*	84–100%
R71–Alluvial barren and debris-covered ice	19	95%*	84–100%
B31–Broadleaf riparian and swamp forest	32	94%*	85–100%
C21–Mountain hemlock, subalpine fir and heather woodland	22	95%*	86–100%
S48–Subalpine heather shrubland	33	94%*	86–100%
S43–Sitka alder shrubland	25	96%*	88–100%
H61–Spreading phlox and common juniper dry meadow	36	97%*	91–100%
B32–Coastal red alder forest	6	100%*	92–100%
C02–Redcedar, Labrador-tea, slough sedge and sphagnum bog	6	100%*	92–100%
C00–Sitka spruce, western hemlock, salal and sword fern forest	10	100%*	95–100%
M90–Ruderal meadow	10	100%*	95–100%
R72–Colluvial barren	13	100%*	96–100%
C12–Silver fir, hemlock and Alaska blueberry forest	16	100%*	97–100%

The class-specific producer's accuracies, obtained from the PCT, are summarized in **Table 16**. The accuracy estimate met the 80% standard for 34 of the 44 map classes. However, for only 13 of them was the 90% confidence interval entirely above the 80% mark. Most of the ten classes that failed to meet the standard were poorly sampled. Only one, C12–SILVER FIR, HEMLOCK AND ALASKA BLUEBERRY FOREST, was conclusively demonstrated to fail to meet the standard. PA could not be estimated for C15–LODGEPOLE PINE AND DOUGLAS-FIR WOODLAND, which was not sampled in the field.

Table 16. Map class-specific producer's accuracy (PA)for each assessed map class, with poorly mapped classes first. Figures are taken from population contingency table. Asterisks indicate true values that are at least 90% confident to lie either fully above or below the 80% accuracy target.

Class code and full name	# of plots identified as	PA estimate	PA 90% conf int
C15–Lodgepole pine and Douglas-fir woodland	0	NA	NA
M90–Ruderal meadow	13	50%	20–80%
C12–Silver fir, hemlock and Alaska blueberry forest	25	59%*	44–75%
B32–Coastal red alder forest	7	69%	26–100%
C14–Silver fir, big huckleberry and beargrass forest	9	68%	27–100%
H60W–Black alpine sedge wetland	5	72%	29–100%
H50W–Lowland marsh and meadow	10	69%	29–100%
C21–Mountain hemlock, subalpine fir and heather woodland	31	67%	53–81%
S40W–Low elevation shrub-dominated wetland	3	80%	37–100%
C05–Western hemlock, Douglas-fir and sword fern forest	34	71%	58–84%
S44–Thimbleberry shrubland, tall forbs and bracken fern	20	71%	54–89%
C00–Sitka spruce, western hemlock, salal and sword fern forest	13	79%	59–98%
S42–Sitka willow riparian shrubland	5	87%	59–100%
C06–Western hemlock, Douglas-fir and salal forest	25	80%	68–93%
H62–Alpine sparse herbaceous vegetation	6	85%	62–100%
C11–Mesic silver fir and western hemlock forest	14	82%	65–99%
H53–Showy sedge and Sitka valerian meadow	34	82%	69–96%
B33–Upland red alder, bigleaf maple and conifer forest	20	83%	65–100%
S47–Successional huckleberry shrubland	31	83%	66–100%
C03–Sitka spruce, western hemlock and wood-sorrel forest	57	85%	75–94%
C01–Redcedar, western hemlock, salal and evergreen huckleberry forest	19	86%	69–100%
C20–Subalpine fir and Sitka valerian forest and woodland	33	85%	72–98%
S49–Alpine heather shrubland	12	89%	70–100%
H51W–Subalpine herbaceous wetland	13	86%	71–100%
H52–Cow parsnip meadow	20	89%	71–100%
S48–Subalpine heather shrubland	41	85%	78–93%
C08–Olympic dry Douglas-fir forest	20	86%	73–99%
H55–Olympic aster scree meadow	14	89%	73–100%
C04–Moist western hemlock, Douglas-fir and foamflower forest	25	87%	76–97%
H54–Moist talus vegetation	2	100%	75–100%
H58–Bedrock balds and sparsely vegetated forest openings	2	100%	75–100%
H56–Subalpine summer-dry grass-forb meadow	34	90%	77–100%
C13–Mountain hemlock, silver fir and Cascade azalea forest	8	94%*	80–100%
S43–Sitka alder shrubland	26	91%*	81–100%
H61–Spreading phlox and common juniper dry meadow	39	91%*	81–100%

Table 16 (continued). Map class-specific producer's accuracy (PA) for each assessed map class, with poorly mapped classes first. Figures are taken from population contingency table. Asterisks indicate true values that are at least 90% confident to lie either fully above or below the 80% accuracy target.

Class code and full name	# of plots identified as	PA estimate	PA 90% conf int
C10–Moist silver fir, western hemlock and foamflower forest	30	93%*	84–100%
B30–Successional gravel bar shrubland	15	97%*	90–100%
R72–Colluvial barren	14	97%*	90–100%
C02–Redcedar, Labrador-tea, slough sedge and sphagnum bog	6	100%*	92–100%
B31–Broadleaf riparian and swamp forest	31	98%*	94–100%
S45–Vine maple shrubland	12	100%*	96–100%
C26–Conifer krummholz and treed cliff	15	100%*	97–100%
R71–Alluvial barren and debris-covered ice	18	100%*	97–100%
S41W–Subalpine willow wetland	23	100%*	98–100%
C24–Olympic subalpine fir and lodgepole pine woodland	25	100%*	98–100%

Classes with accuracies less than the 80% target should be considered as candidates for merging with other classes (Lea and Curtis 2010). To assist with this task, **Table 17** lists these classes as well as the classes with which each is confused.

Table 17. Significantly confused map classesand the classes with which they are confused. For classes with user's or producer's accuracy less than 80%, all classes with which confusion exists are given with the proportion of the reached inference area (RIA) affected by confusion in either direction between the pair.

Class code and abbreviated name	Minimum (UA, PA)	Confused with (proportion of RIA affected)
C15–Lodgepole pine & Doug-fir	0%	C08–Olympic dry Doug-fir (0.17%) C06–W hemlock & salal (0.09%)
M90–Ruderal meadow	50%	B30–Successional gravel bar (0.07%) H50W–Lowland marsh & meadow (0.03%) S44–Thimbleberry, forbs & bracken fern (0.02%)
C14–Silver fir & big huckleberry	50%	C12–Silver fir & Alaska blueberry (0.38%) C13–Mtn hemlock & Cascade azalea (0.36%) C11–Mesic silver fir & w hemlock (0.19%) C10–Moist silver fir & foamflower (0.10%) C05–W hemlock & sword fern (0.10%)
H54–Moist talus vegetation	50%	R72–Colluvial barren (0.04%) S48–Subalpine heather (0.04%)
C13–Mtn hemlock & Cascade azalea	58%	C21–Mtn hemlock & heather (0.71%) C20–Subalp fir & Sitka valerian (0.50%) C12–Silver fir & Alaska blueberry (0.36%) C14–Silver fir & big huckleberry (0.36%)

Table 17 (continued). Significantly confused map classes and the classes with which they are confused. For classes with user's or producer's accuracy less than 80%, all classes with which confusion exists are given with the proportion of the reached inference area (RIA) affected by confusion in either direction between the pair.

Class code and abbreviated name	Minimum (UA, PA)	Confused with (proportion of RIA affected)
C12–Silver fir & Alaska blueberry	59%	C11–Mesic silver fir & w hemlock (0.85%) C14–Silver fir & big huckleberry (0.38%) C13–Mtn hemlock & Cascade azalea (0.36%) C10–Moist silver fir & foamflower (0.23%)
C11–Mesic silver fir & w hemlock	67%	C05–W hemlock & sword fern (1.00%) C12–Silver fir & Alaska blueberry (0.85%) C14–Silver fir & big huckleberry (0.19%)
S42–Sitka willow riparian	67%	S43–Sitka alder (0.10%) S45–Vine maple (0.03%)
C21–Mtn hemlock & heather	67%	C20–Subalp fir & Sitka valerian (1.03%) C13–Mtn hemlock & Cascade azalea (0.71%) C24–Olympic subalp fir & lodgepole pine (0.07%) S41W–Subalpine willow wetland (0.02%)
B32–Coastal red alder	69%	C03–Sitka spruce & wood-sorrel (0.18%)
H50W–Lowland marsh & meadow	69%	R71–Alluvial barren (0.11%) B30–Successional gravel bar (0.03%) S40W–Low elevation shrub wetland (0.03%) M90–Ruderal meadow (0.03%)
C05–W hemlock & sword fern	71%	C06–W hemlock & salal (2.09%) C11–Mesic silver fir & w hemlock (1.00%) C04–Moist w hemlock & foamflower (0.57%) C10–Moist silver fir & foamflower (0.23%) C08–Olympic dry Doug-fir (0.22%) C14–Silver fir & big huckleberry (0.10%)
H51W–Subalpine herbaceous wetland	71%	S47–Successional huckleberry (0.05%) S41W–Subalpine willow wetland (0.04%) H60W–Black alpine sedge (0.03%) H52–Cow parsnip (0.03%) S48–Subalpine heather (0.03%)
S44–Thimbleberry, forbs & bracken fern	71%	H52–Cow parsnip (0.06%) S43–Sitka alder (0.04%) M90–Ruderal meadow (0.02%) S47–Successional huckleberry (0.01%)
H60W–Black alpine sedge	72%	H51W–Subalpine herbaceous wetland (0.03%) H53–Showy sedge & Sitka valerian (0.02%)
C06–W hemlock & salal	73%	C05–W hemlock & sword fern (2.09%) C03–Sitka spruce & wood-sorrel (0.50%) C08–Olympic dry Doug-fir (0.25%) C15–Lodgepole pine & Doug-fir (0.09%)
C00–Sitka spruce, salal & sword fern	79%	C03–Sitka spruce & wood-sorrel (0.37%) C01–Redcedar, salal & evergreen huckleberry (0.13%)

Table 17 (continued). Significantly confused map classes and the classes with which they are confused. For classes with user's or producer's accuracy less than 80%, all classes with which confusion exists are given with the proportion of the reached inference area (RIA) affected by confusion in either direction between the pair.

Class code and abbreviated name	Minimum (UA, PA)	Confused with (proportion of RIA affected)
H52–Cow parsnip	79%	S44–Thimbleberry, forbs & bracken fern (0.06%), H51W–Subalpine herbaceous wetland (0.03%) H53–Showy sedge & Sitka valerian (0.01%)

Sample and population contingency tables were also constructed at the lifeform/land-use level by lumping each map class into a category based on its dominant vegetation; the results are in INR (2021a) and are summarized below in **Table 18**. All AA plots that successfully passed through the QC process were used, including those of the nine map classes excluded from the map class level analysis (these fell into the last five categories in the table).

Table 18. Accuracy of map aggregated to lifeform/land-use level Figures taken from population contingency table. Asterisks indicate true values that are at least 90% confident to lie either fully above or below the 80% accuracy target.

Lifeform	# of plots mapped as	UA estimate	UA 90% conf interval	# of plots identified as	PA estimate	PA 90% conf interval
Broadleaf tree	72	96%*	91–100%	73	95%*	90–100%
Conifer	390	99%*	98–100%	389	99%*	99–100%
Tall shrub	44	98%*	93–100%	43	100%*	99–100%
Shrubland	123	91%*	86–96%	130	80%	72–89%
Herbaceous	188	90%*	86–94%	179	95%*	92–98%
Water	64	97%*	93–100%	62	100%*	99–100%
Rock	76	99%*	96–100%	78	97%*	93–100%
Snow & ice	20	100%*	98–100%	20	100%*	98–100%
Developed	1	100%	50–100%	1	100%	50–100%
Other disturbed	13	100%*	96–100%	16	64%	32–96%

The user's and producer's accuracy estimates exceeded the 80% standard for all natural land-use categories, all at 90% confidence except for the **SHRUBLAND** producer's accuracy. This failure was largely due to three errors: mapping S48–SUBALPINE HEATHER SHRUBLAND as H53–SHOWY SEDGE AND SITKA VALERIAN MEADOW, mapping S44–THIMBLEBERRY SHRUBLAND, TALL FORBS AND BRACKEN FERN as H52–COW PARSNIP MEADOW, and mapping S47–SUCCESSIONAL HUCKLEBERRY SHRUBLAND as H55–OLYMPIC ASTER SCREE MEADOW. The producer's accuracy estimate for **OTHER DISTURBED**—which included M90–RUDERAL MEADOW, M92–BURNED WITH UNCERTAIN VEGETATION and M93–TIMBERLAND WITH UNCERTAIN VEGETATION—was lower than 80%, though the sample size was small and it was not conclusively demonstrated to fail to meet the standard via the 90%

confidence interval. The poor result was due to mapping of three ruderal meadow plots as other vegetation types common in disturbed areas.

3.7. Discussion

The accuracy assessment phase field observations are the reference by which the map's accuracy is measured, but these observations are not infallible. Key decisions in the field regarding the extent of the vegetation types perceived, the locations of boundaries between them, and the cover occupied by the species present within each may vary between observers. The vegetation classification itself (Nielsen et al. 2021c, Nielsen and Brunner 2021) is a somewhat subjective entity, with few hard rules for discriminating classes other than the weight of statistical evidence from ocular data, which are incomplete for many AA plots. Observers may also disagree about the degree to which a text-based map class description matches a vegetation patch in the field. To borrow a term from taxonomy, the circumscriptions of the map classes and mapping associations may not be consistently understood and applied. In some cases, the accuracy assessment plot quality control process will have compensated for these inconsistencies; in other cases, not.

The failure to meet inference area goals for most classes and the geographic bias toward areas of the park which we were able to reach limit the confidence we can attach to many of our conclusions. In these cases, Lea and Curtis (2010) warn that "extending the results of the thematic accuracy assessment from the inference area to the rest of the project must be justified by assumptions, rather than by statistical inference." In the following discussion, we have supplemented the AA analysis with photo-interpretation and consideration of context, in an attempt to provide additional evidence and to make these assumptions as transparent as possible.

3.7.1. Undersampled map classes

Based on their mapped area, the assessed classes each require 30 samples of mapped occurrences. As documented in **Table 14**, this was only achieved for ten of the 45 classes. Several causes for this failure are described below.

Fourteen of the 35 undersampled classes were rare (totaling less than 100 mapped hectares) in the reached inference area. These classes were typically concentrated in one or two parts of the RIA and even if the sampling goals had been achieved, autocorrelation amongst these plots in both floristics and mapping tendencies would have likely made their application to the full project area questionable.

Five other classes (C01–REDCEDAR, WESTERN HEMLOCK, SALAL AND EVERGREEN HUCKLEBERRY FOREST, C26–CONIFER KRUMMHOLZ AND TREED CLIFF, H55–OLYMPIC ASTER SCREE MEADOW, H62–ALPINE SPARSE HERBACEOUS VEGETATION, and R72–COLLUVIAL BARREN) are very often located in inaccessible areas, either due to high elevations and steep slopes or to impenetrable vegetation. It is likely that field efficiency or safety concerns often resulted in crews skipping plots from these classes in order to maximize overall plot collection or stay on their itinerary.

Of the remaining 16 undersampled classes, only three had fewer than 15 mapped occurrences. The extreme undersampling of these three classes (C00–SITKA SPRUCE, WESTERN HEMLOCK, SALAL AND

SWORD FERN FOREST, C13–MOUNTAIN HEMLOCK, SILVER FIR AND CASCADE AZALEA FOREST, and H51W–SUBALPINE HERBACEOUS WETLAND), in addition to the failure to meet the standard for several of the other remaining classes, was likely due to changes in the map classification made between the time of AA sampling in 2013 and production of the final maps in 2019–20. For example, during the intervening time period, we addressed substantial confusion between H51W–SUBALPINE HERBACEOUS WETLAND and H53–SHOWY SEDGE AND SITKA VALERIAN MEADOW by carving a new map class, S41W–SUBALPINE WILLOW WETLAND, out of plots with substantial *Salix commutata* presence that had floristics and mapping tendencies intermediate to the two original classes.

Many other changes were made to address challenges of mapping and field identification. While the goals of these updates were achieved, they had the by-product of reducing the number of AA plots available for assessing the final classes. An ideally executed project would have deferred AA fieldwork until completion of the final map—or at least map classification—but project management concerns took precedence here. For many map classes, the sample sizes are too small to confidently assess whether the 80% accuracy standard was achieved, as reflected in the wide confidence intervals seen in **Table 15** and **Table 16**. The small sample sizes should be kept in mind when considering the mapping error rates discussed below.

3.7.2. Map classes failing to meet accuracy standards

A list of the map classes failing to meet accuracy standards, the classes they are most confused with, and a possible corrective action that could be taken (if any) are shown in **Table 19**. Since every area of the map must be labeled, the only corrective action we consider is that of merging confused classes. This is likely to result in overall improvements only if the classes to be merged are confused primarily with each other. Otherwise, any poorly mapping area will simply get attributed into a different bin, perhaps bringing a different class below the accuracy target. This situation applies to several of the map classes with the most egregious errors. For example, based on the three samples obtained, C15–LODGEPOLE PINE AND DOUGLAS-FIR WOODLAND appears to be most confused with C08–OLYMPIC DRY DOUGLAS-FIR FOREST and C06–WESTERN HEMLOCK, DOUGLAS-FIR AND SALAL FOREST. However, each of these classes is primarily confused with yet other classes. Merging C15 with either of them would merely reapportion the error. We first review the apparent mapping errors for which merging classes does not appear to be an option.

Table 19. Map classes failing to meet accuracy standardsor confused with those classes. '+' indicates accuracy estimates of 80% or higher; asterisks indicate accuracy less than 80% at 90% confidence. The classes accounting for the most mismapped area are listed under "confusion with," along with the fraction contributed to the total mismapped area in parentheses. A possible corrective action is noted for each.

Class code and abbreviated name	UA %	PA %	Confusion with	Corrective action
C00–Sitka spruce, salal & sword fern	+	72	C03 (74%), C01 (26%)	none, no reciprocity
C03–Sitka spruce & wood-sorrel	+	+	C04 (31%), B31 (18%)	–
C05–W hemlock & sword fern	74	71	C06 (50%), C11 (24%)	consider merge with C06
C06–W hemlock & salal	73	+	C05 (71%)	consider merge with C05
C08–Olympic dry Doug-fir	+	+	B33 (43%), C06 (22%)	–
C11–Mesic silver fir & w hemlock	67	+	C05 (49%), C12 (42%)	consider merge with C12
C12–Silver fir & Alaska blueberry	+	50*	C11 (47%), C14 (21%)	consider merge with C11
C13–Mtn hemlock & Cascade azalea	58	+	C21 (37%), C20 (26%)	consider merge with C21 and C20
C14–Silver fir & big huckleberry	50*	68	C12 (34%), C13 (32%)	none, no reciprocity
C15–Lodgepole pine & Doug-fir	0*	NA	C08 (67%), C06 (33%)	none, no reciprocity
C20–Subalp fir & Sitka valerian	+	+	C21 (59%), C13 (29%)	consider merge with C13 and C21
C21–Mtn hemlock & heather	+	59	C20 (56%), C13 (39%)	consider merge with C13 and C20
B30–Successional gravel bar	+	+	M90 (72%), H50W (28%)	consider merge with M90
B32–Coastal red alder	+	68	C03 (100%)	none, no reciprocity
S42–Sitka willow riparian	67	+	S43 (77%)	consider merge with S43
S43–Sitka alder	+	+	S42 (71%), S44 (29%)	consider merge with S42
S44–Thimbleberry, forbs & bracken fern	+	71	H52 (43%), S43 (31%)	consider merge with H52
S47–Successional huckleberry	+	+	H55 (35%), H51W (34%)	–
H50W–Lowland marsh & meadow	75	69	R71 (58%)	none, can't merge with abiotic
H51W–Subalpine herbaceous wetland	71	+	S47 (26%), S41W (21%)	none, no reciprocity
H52–Cow parsnip	79	+	S44 (55%), H51W (32%)	consider merge with S44
H54–Moist talus vegetation	50	+	R72 (50%), S48 (50%)	none, can't merge with abiotic
H60W–Black alpine sedge	+	71	H51W (61%), H53 (39%)	none, no reciprocity
M90–Ruderal meadow	+	50	B30 (60%), H50W (23%)	consider merge with B30

Map classes for which merging is not a viable option

No remedy is possible for the following apparent mapping errors, due to non-reciprocity of errors within the confused classes. The classes are considered in order of decreasing severity. Asterisks indicate estimates that are 90% confident to lie either above or below the 80% accuracy target; all other estimates given are not statistically significant with respect to the target. Recommendations to NPS are given in boldface.

C15–LODGEPOLE PINE AND DOUGLAS-FIR WOODLAND (UA 0%*, PA NA) is likely over-mapped on dry slopes with canopy openings on the east side, but the sample size is very small. One patch of C08–OLYMPIC DRY DOUGLAS-FIR FOREST had two plots that mapped as C15, and in another place C15 was mapped where C06–WESTERN HEMLOCK, DOUGLAS-FIR AND SALAL FOREST was found on the ground. The AA sample included none of the known C15 patches in the Elwha, which appear by PI to be well-mapped. ***We recommend NPS make additional effort to verify the map, both above the Elwha and on the east side.***

C14–SILVER FIR, BIG HUCKLEBERRY AND BEARGRASS FOREST (UA 50%*, PA 68%) appears somewhat over-mapped, *mapping over*[71] C12–SILVER FIR, HEMLOCK AND ALASKA BLUEBERRY FOREST in east-side drainages and the northwest corner, and over C11–MESIC SILVER FIR AND WESTERN HEMLOCK FOREST on the upper Gray Wolf River. It is likely that C14 should be more highly restricted to valley walls than currently mapped, although the plots had lower landscape positions than is typical for C14. This class occurs in settings intermediate to a range of montane and upper montane classes, several of which have depauperate tendencies. Field identification can be inconclusive, and the class tends to absorb many of the mapping errors in these intermediate zones.

H54–MOIST TALUS VEGETATION (UA 50%, PA 100%) appears to map over R72–COLLUVIAL BARREN and H62–ALPINE SPARSE HERBACEOUS VEGETATION occasionally. Sample sizes were extremely low, but inspection of the map against 2015 NAIP imagery seems to confirm confusion with the former, at any rate, although the habitat itself appears to be appropriate for H54. Merging the classes is not an option, as there is no doubt that similar errors exist between other sparsely vegetated classes and more sparsely vegetated areas mapped as one of the abiotic classes. This error is related to the need for some arbitrary vascular vegetation cutoff for distinguishing between vegetated and unvegetated classes. 10% was chosen as it was already in use in the NVC and appeared to be a generally mappable threshold. The use of unvegetated classes is a coping mechanism that is necessary to address the difficulties of either mapping or classifying extremely sparse vegetation, and there will always be both mapping and classification uncertainty in these areas. ***We recommend that NPS pursue some additional validation to establish whether indeed H54 and perhaps other sparse vegetation types are over-mapped, but also to regard these sparse types from a habitat-driven point of view rather than always bearing a direct relationship to the amount of vascular cover present.***

B32–COASTAL RED ALDER FOREST (UA 100%*, PA 69%) had seven field-identified plots, one of which mapped as C03–SITKA SPRUCE, WESTERN HEMLOCK AND WOOD-SORREL FOREST. This translated into a substantial producer's error since C03 is widely distributed. The error occurred at a plot that was heavily disturbed by logging and road construction. There was in fact a small patch of conifers present, but the overall species composition was a better match to B32, which can contain conifers. In general, classification accuracy, and hence map accuracy, will be lower in disturbed areas, since the NVC typically does not explicitly address anthropogenic disturbance. There are many areas of the map, particularly toward the coast, where broadleaf and conifer forest intergrade on a fine scale;

[71] In this terminology, "X *maps over* Y" means that class X was repeatedly mapped in locations where class Y was discovered on the ground. "Y *is mapped over by* X" would be an equivalent formulation.

errors at our mapping resolution will invariably occur. ***Within highly spatially intermeshed vegetation types, and in other floristically intermediate areas, we recommend that NPS regard mosaics of small mapped forest patches as indicating the proportions in which several vegetation types may be represented locally, rather than always seeking a fine-scale spatial correspondence.***

H50W–LOWLAND MARSH AND MEADOW (UA 75%, PA 69%) appears occasionally over-mapped along rivers and streams and in canopy gaps on floodplains, where M90–RUDERAL MEADOW, B30–SUCCESSIONAL GRAVEL BAR SHRUBLAND or S40W–LOW ELEVATION SHRUB-DOMINATED WETLAND might be a better option. These classes were all rather poorly trained and assessed, due to limited sampling opportunities. ***We recommend that NPS undertake a field sampling and photo-interpretation effort to determine more precisely the degree of H50W over-mapping and to gradually refine the map while doing so.***

H51W–SUBALPINE HERBACEOUS WETLAND (UA 71%, PA 86%) appears to be over-mapped in some environments: it maps over one plot each of H52–COW PARSNIP MEADOW, H60W–BLACK ALPINE SEDGE WETLAND, S47–SUCCESSIONAL HUCKLEBERRY SHRUBLAND, and S48–SUBALPINE HEATHER SHRUBLAND. Conversely, it is mapped over by S41W–SUBALPINE WILLOW WETLAND on two adjacent plots on the east side. H51W often occurs in highly heterogeneous areas where some degree of error is inevitable. ***As above, within highly spatially intermeshed vegetation types, and in other floristically intermediate areas, we recommend that NPS regard mosaics of small mapped patches as indicating the proportions in which several vegetation types may be represented locally, rather than always seeking a fine-scale spatial correspondence.***

H60W–BLACK ALPINE SEDGE WETLAND (UA 80%, PA 72%) had six field-identified plots, one of which mapped as H51W. This resulted in a large producer error since H51W occupies a much larger area of the map. The error occurred in a narrow strip of meadow subject to cold air drainage and dominated by *Carex nigricans* and *Luetkea pectinata*. H60W is a challenging class for mapping due to its small patch size, although it is readily identified in the field. Because the background upland vegetation in which it is embedded varies, it cannot be effectively merged with any other class.

C00–SITKA SPRUCE, WESTERN HEMLOCK, SALAL AND SWORD FERN FOREST (UA 100%*, PA 79%) is mapped over by C03–SITKA SPRUCE, WESTERN HEMLOCK AND WOOD-SORREL FOREST on two plots very near water, once by Ozette Lake, and once near the lowest reach of the Queets River in the park interior. Distance to water appears to have been one of the significant breakpoints our models used for distinguishing these classes, with C00 generally found in more upland settings than C03 within the overlap of their geographic ranges. That conclusion may have been unwarranted based on these plots. There are other mapping errors in this area, which is geographically dissimilar to any other part of the park. AA data suggest that both C00 and C01–REDCEDAR, WESTERN HEMLOCK, SALAL AND EVERGREEN HUCKLEBERRY FOREST extend farther from the coast here than we have mapped. ***We recommend that NPS assess the lowest portion of the Queets River in the park interior and consider redrawing the boundaries of C00, C01 and C03 here by hand.***

Map classes for which merging may be a viable option

There are six groups of map classes that could conceivably be aggregated for improved accuracy, based on this analysis. They are considered below in order of decreasing severity of the apparent mapping issue they would address. Asterisks indicate estimates that are 90% confident to lie either above or below the 80% accuracy target; all other estimates given are not statistically significant with respect to the target. Recommendations to NPS are given in boldface.

C12–SILVER FIR, HEMLOCK AND ALASKA BLUEBERRY FOREST (UA 100%*, PA 59%*) and C11–MESIC SILVER FIR AND WESTERN HEMLOCK FOREST (UA 67%, PA 82%) could be merged. The combined class C11+C12 would have UA 94%* and PA 80%. Confusion between C11 and C12 is documented at only three plots, two of which are immediately adjacent to each other in a small patch in the northwestern corner of the park where half of the C12 omission error was found. This raises doubts about whether the C12 PA is well-estimated and whether confusion with C11 is really the largest issue. The other incorrectly mapped plot was mapped as a small patch of C11 surrounded by mapped C12, in a transitional area of the map. Furthermore, despite the fact that C11 was the most abundantly mapped class in the park, with nearly twice the area of the second most abundant class, it was also the least well-represented analyzed map class in the inference area, with only 1.6% of its mapped area falling in the RIA (C12 was the second most poorly represented, with only 2.4% of its area in the RIA). Both these classes are primarily mapped on valley walls in the western portion of the park, where there are virtually no trails. Given the poor and unrepresentative sampling available, the abundance of the classes, and the fact that each is significantly confused with others, merging them appears unwarranted and more likely to damage the map than to solve any real issues. ***We recommend that NPS conduct additional investigation into the mapping of the west-side valley walls; however, we believe C11 and C12 are best kept as distinct map classes.***

M90–RUDERAL MEADOW (UA 100%*, PA 50%*) and B30–SUCCESSIONAL GRAVEL BAR SHRUBLAND (UA 93%, PA 97%*) could be merged. The combined class B30+M90 would have UA 100%* and PA 94%*. While both M90 and B30 are found near rivers in major valleys, their similarity ends there. In fact, the confusion here is based on a single plot in a weed patch at Enchanted Valley. M90 had not been documented previously in that area, with its training data concentrated in the lower Queets and Elwha Valleys. It does appear to be the best label for the location, but the classes should not be merged on this basis. In fact, M90 is confused with one plot each of two other map classes: one also in Enchanted Valley, and one at Cape Alava campground. It seems likely that the extreme over-representation of anthropogenic disturbance in the inference area is responsible for the apparently low M90 PA. We believe that across the full park the effective M90 PA is likely to be well over 80%. ***We recommend that NPS hand-correct what we believe to be a limited number of additional M90 patches that were not mapped as such, concentrating on areas of known anthropogenic disturbance.***

C13–MOUNTAIN HEMLOCK, SILVER FIR AND CASCADE AZALEA FOREST (UA 58%, PA 94%*), C21–MOUNTAIN HEMLOCK, SUBALPINE FIR AND HEATHER WOODLAND (UA 95%*, PA 67%), and C20–SUBALPINE FIR AND SITKA VALERIAN FOREST AND WOODLAND (UA 80%, PA 85%) could be merged. The combined class C13+C20+C21 would have UA 97%* and PA 98%*. Merging the three classes

would create a single class representing all mesic subalpine forests and woodlands and spanning considerable floristic variability; C13 was both poorly represented in the inference area (with just 4.1% of its mapped area within the RIA) and poorly sampled (with only 12 mapped plots assessed). This does not provide strong evidence that it should be merged with the other two classes. Although C20 and C21 were also poorly represented (with 6.5% and 4.2% of their mapped area within the RIA), there was considerably more evidence of their legitimate mismapping: there was confusion on seven plots, six of them mapped as C20 but identified in the field as C21 (and mostly well-dispersed). Unfortunately, the floristics of these classes are not very similar: C20 has greater affinity with C24–OLYMPIC SUBALPINE FIR & LODGEPOLE PINE and C13, and C21 has greater affinity with C13 also. Merging C20 and C21 without including C13 would create significant difficulties in distinguishing C13 from the merged C20+C21. It appears from this evidence that it would be preferable to merge all three classes instead, despite the lack of strong evidence that C13 is poorly mapped and the fact that the C20 currently meets the accuracy standards. ***We recommend the classes be kept separate, with the caveat that C21 appears to be under-mapped in the park's northeast section, where it may be erroneously labeled as C20. We base this recommendation on the fact that the confusion appears confined to a narrow set of conditions (tree islands with*** *Tsuga mertensiana* ***dominant in a landscape mostly dominated by*** *Abies lasiocarpa****) in fairly narrow geographic portion of the full ranges of the three involved map classes. Furthermore, C21 is correctly mapped in many of the adjacent areas with sapling-sized trees; the error appears to be more of degree than of kind. However, we request NPS to evaluate and advise.***

S42–SITKA WILLOW RIPARIAN SHRUBLAND (UA 67%, PA 87%) and S43–SITKA ALDER SHRUBLAND (96%*, 91%*) could be merged. The combined class S42+S43 would have UA 97%* and PA 98%*. The S42 UA was estimated from only six mapped plots, with the two incorrectly mapped plots only 100 meters apart. Both were surveyed by a less-experienced member of the team, working alone; no floristic information was provided apart from the presence of dense *Alnus viridis*. The setting of these plots—low on a south-facing slope just above the Dosewallips River—is appropriate for S42, and only a small amount of *Salix sitchensis* would have been needed to make S42 an appropriate call. In addition, *Heracleum maximum* (an indicator for S42 vs. S43) is documented just below the plot. Furthermore, S42 was a class created after the AA sampling was performed, so was not a labeling option at the time of fieldwork. While we consider these plots to have been mapped incorrectly, we don't believe they provide evidence for merging these classes. ***We recommend the classes be kept separate.***

C05–WESTERN HEMLOCK, DOUGLAS-FIR AND SWORD FERN FOREST (UA 74%, PA 71%) and C06–WESTERN HEMLOCK, DOUGLAS-FIR AND SALAL FOREST (UA 73%, PA 80%) could be merged. The combined class C05+C06 would have UA 88%* and PA 90%*. The current accuracies are not far below 80% and the incorrectly labeled plots mostly lay within meters of the mapped transition between the two classes. It appears that the RIA happened to bias sampling within the transitional zone. The locations of the incorrect plots do not provide evidence that the broad zones within which the two classes are mapped are in error. This is particularly true for C06, which is widely distributed (mapped on 30,128 ha) and largely inaccessible from trails (4.8% within RIA). ***Given the lack of convincing evidence that the classes are poorly mapped, and their reasonably high accuracies as***

mapped, we recommend they be kept separate, but that NPS conduct additional investigation into the mapping of the valley walls above the Elwha River and Maiden and Morse Creeks, where a substantial fraction of the C06 mapped distribution occurs.

S44–THIMBLEBERRY SHRUBLAND, TALL FORBS AND BRACKEN FERN (UA 93%, PA 71%) and H52–COW PARSNIP MEADOW (UA 79%, PA 89%) could be merged. The combined class S44+H52 would have UA 95%* and PA 87%. Despite one being shrub-dominated and the other herbaceous, they are often spatially intermingled. However, both classes show significant confusion with other types: S44 with S43–SITKA ALDER SHRUBLAND, and H52 with H51W–SUBALPINE HERBACEOUS WETLAND. Confusion between S44 and H52 was documented on four east-side plots, all of which were mapped as H52 but identified in the field as S44. Three of these plots—all within 400 meters of each other along Cameron Creek—were dominated by creeping snowberry, *Symphoricarpos mollis*, which rises little if at all above the surrounding herbaceous vegetation. The other plot was dominated by *Rubus nutkanus* and should have been mappable as S44. This error may indicate a serious under-mapping of S44 on the east side that may have resulted from bias in the training data. ***Based on their distinct lifeforms, as well as their confusion with different types of the same lifeform, we recommend these classes be kept separate—with the caveat that some areas dominated by*** *S. mollis* ***with herbaceous vegetation will inevitably map as H52. However, we also recommend that additional sampling, beginning in the area around Dose Meadows, be undertaken to ascertain the degree to which S44 was under-mapped in this area and on the east side in general. If further sampling indicates a great enough problem, NPS may want to merge S44 and H52 in the future.***

3.7.3. Other known mapping issues

We observed several other mapping issues that will likely be noticed from time to time. ***Further investigation of these issues by NPS may be warranted.***

H53–SHOWY SEDGE AND SITKA VALERIAN MEADOW may map over lush S48–SUBALPINE HEATHER SHRUBLAND where *Vaccinium deliciosum* is present in large quantities, especially if there is substantial *Carex spectabilis* and/or *Lupinus latifolius*. In these circumstances, the *Phyllodoce empetriformis* and *Cassiope mertensiana* may be all but invisible in imagery. This likely simply reflects a difficulty posed by the emphasis on lifeform in the classification, as in many cases the difference between these classes is more of quantity than of quality.

H52–COW PARSNIP MEADOW was discussed above pertaining to its confusion with S44–THIMBLEBERRY SHRUBLAND, TALL FORBS AND BRACKEN FERN, but there is also some indication of confusion with H53–SHOWY SEDGE AND SITKA VALERIAN MEADOW. Making consistent field calls between H52 and H53 requires comprehensive floristic data, because H52 is recognized not by the absence of subalpine species such as *Carex spectabilis*, *Veratrum viride*, and *Valeriana sitchensis*, but by the presence of any of a wide range of montane meadow species which may be difficult for crews to identify.

S49–ALPINE HEATHER SHRUBLAND was mapped at two plots that were field-identified as H62–ALPINE SPARSE HERBACEOUS VEGETATION. It is likely this occurs in other places as well. S49 is generally a difficult class as it varies widely in density and appearance.

H56–SUBALPINE SUMMER-DRY GRASS-FORB MEADOW was mapped at two plots that were field-identified as H61–SPREADING PHLOX AND COMMON JUNIPER DRY MEADOW. As these plots were in the heart of the geographic range of both these classes in the northeastern corner, it is likely this error reflects some over-mapping of H56.

4. Vegetation of Olympic National Park

4.1. Vegetation map

The vegetation map (**Figure 20**, and a higher-resolution version at Nielsen et al. 2021d) illustrates the distribution of the 54 map classes across the park and surrounding buffer. The map contains nearly 700 million pixels aggregated into patches of no less than nine 3-meter pixels (81 m^2). The estimated area of each class, based on its mapped area modified by the correction factor from the AA population contingency table, is shown in **Table 20**. The map classes vary widely in abundance, with most of them limited in extent. The most common ten classes—all of which are conifer-dominated types—collectively occupy 70% of the park, while 32 classes cover less than 1% each. M93–TIMBERLAND WITH UNCERTAIN VEGETATION is absent in the park, though it is present immediately adjacent in several areas; it is not shown in the table.

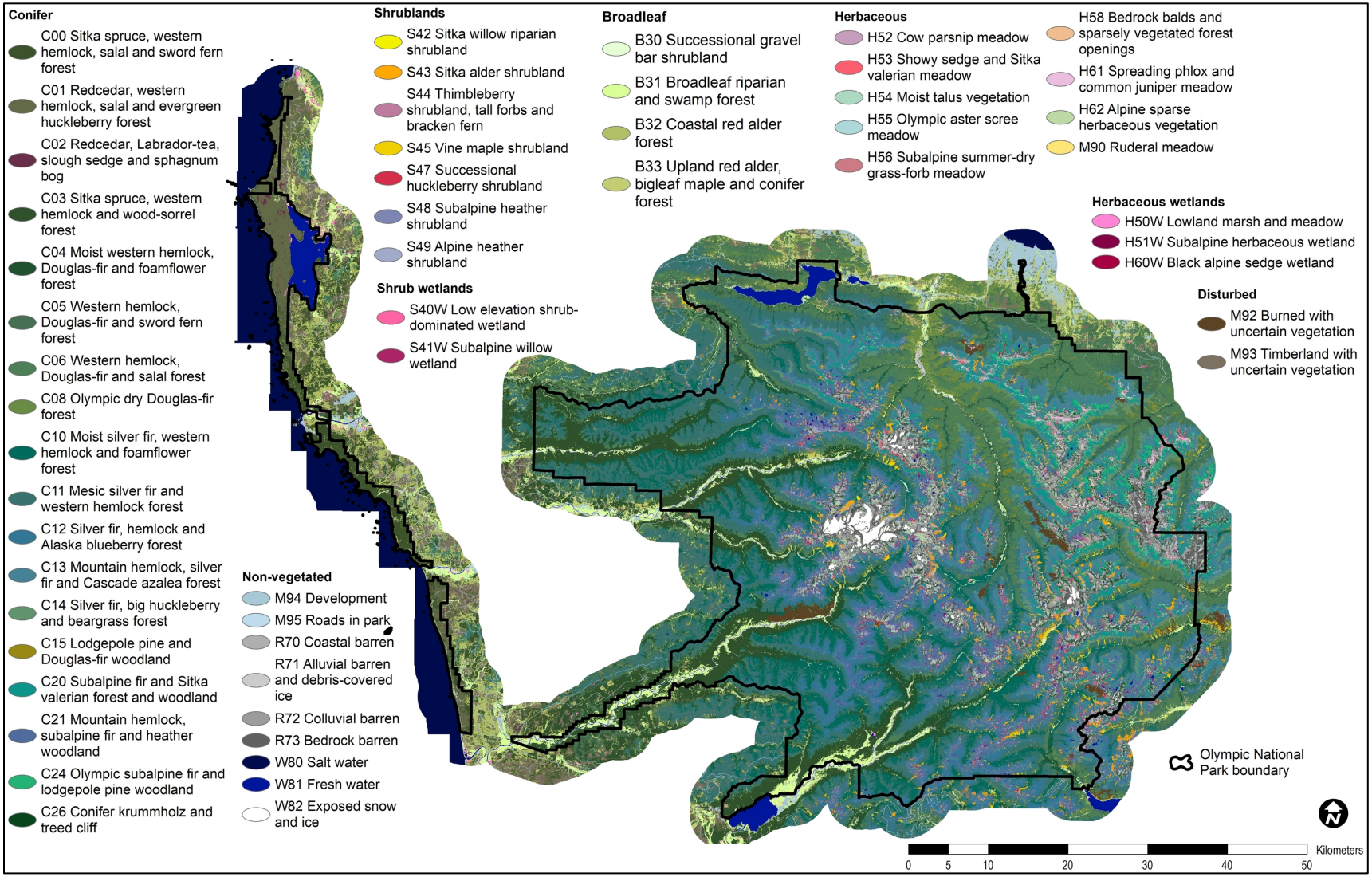


Figure 20. Vegetation map of Olympic National Park.

Table 20. Map class estimated area and proportion of park listed by area.

Class code and full name	Est area in park (ha)	Est area in park (acres)	Proportion of park (%)
C11–Mesic silver fir and western hemlock forest	53,185	131,422	14.38
C12–Silver fir, hemlock and Alaska blueberry forest	34,794	85,979	9.41
C10–Moist silver fir, western hemlock and foamflower forest	29,797	73,630	8.06
C06–Western hemlock, Douglas-fir and salal forest	26,561	65,633	7.18
C21–Mountain hemlock, subalpine fir and heather woodland	25,019	61,822	6.76
C03–Sitka spruce, western hemlock and wood-sorrel forest	21,902	54,121	5.92
C04–Moist western hemlock, Douglas-fir and foamflower forest	19,873	49,108	5.37
C20–Subalpine fir and Sitka valerian forest and woodland	17,292	42,730	4.67
C05–Western hemlock, Douglas-fir and sword fern forest	15,448	38,173	4.18
C13–Mountain hemlock, silver fir and Cascade azalea forest	14,149	34,962	3.82
R72–Colluvial barren	9,824	24,276	2.66
R73–Bedrock barren	8,251	20,390	2.23
C08–Olympic dry Douglas-fir forest	6,927	17,117	1.87
C01–Redcedar, western hemlock, salal and evergreen huckleberry forest	6,874	16,986	1.86
W81–Fresh water	6,666	16,472	1.80
C14–Silver fir, big huckleberry and beargrass forest	6,503	16,068	1.76
C00–Sitka spruce, western hemlock, salal and sword fern forest	6,457	15,955	1.75
B31–Broadleaf riparian and swamp forest	5,795	14,319	1.57
C26–Conifer krummholz and treed cliff	5,390	13,319	1.46
S43–Sitka alder shrubland	4,836	11,949	1.31
C24–Olympic subalpine fir and lodgepole pine woodland	4,083	10,090	1.10
S48–Subalpine heather shrubland	3,563	8,804	0.96
H55–Olympic aster scree meadow	3,300	8,155	0.89
B33–Upland red alder, bigleaf maple and conifer forest	2,836	7,009	0.77
M92–Burned with uncertain vegetation	2,827	6,986	0.76
W82–Exposed snow and ice	2,706	6,686	0.73
H62–Alpine sparse herbaceous vegetation	2,671	6,600	0.72
H61–Spreading phlox and common juniper dry meadow	2,560	6,326	0.69
R71–Alluvial barren and debris-covered ice	2,487	6,146	0.67
H53–Showy sedge and Sitka valerian meadow	2,129	5,261	0.58
S47–Successional huckleberry shrubland	1,797	4,440	0.49
S44–Thimbleberry shrubland, tall forbs and bracken fern	1,349	3,334	0.36
S49–Alpine heather shrubland	1,337	3,303	0.36
B32–Coastal red alder forest	1,157	2,860	0.31
B30–Successional gravel bar shrubland	919	2,270	0.25

Table 20 (continued). Map class estimated area and proportion of park, listed by area.

Class code and full name	Est area in park (ha)	Est area in park (acres)	Proportion of park (%)
W80–Salt water	867	2,142	0.23
H51W–Subalpine herbaceous wetland	825	2,040	0.22
S45–Vine maple shrubland	769	1,899	0.21
S41W–Subalpine willow wetland	613	1,515	0.17
H54–Moist talus vegetation	592	1,464	0.16
M95–Roads in park	571	1,411	0.15
H56–Subalpine summer-dry grass-forb meadow	558	1,379	0.15
R70–Coastal barren	556	1,374	0.15
C15–Lodgepole pine and Douglas-fir woodland	555	1,371	0.15
H58–Bedrock balds and sparsely vegetated forest openings	518	1,281	0.14
H52–Cow parsnip meadow	448	1,107	0.12
S42–Sitka willow riparian shrubland	416	1,029	0.11
H50W–Lowland marsh and meadow	416	1,027	0.11
H60W–Black alpine sedge wetland	308	761	0.08
C02–Redcedar, Labrador-tea, slough sedge and sphagnum bog	210	520	0.06
S40W–Low elevation shrub-dominated wetland	147	363	0.04
M90–Ruderal meadow	143	353	0.04
M94–Development	111	275	0.03

4.2. Vegetation overview

The map classes can be broadly broken into ten groups based on their dominant lifeform and land-use characteristics: (a) conifer-dominated, (b) broadleaf tree-dominated, (c) tall shrublands, (d) shrublands and dwarf-shrublands, (e) herbaceous vegetation, (f) rock barrens, (g) exposed snow and ice, (h) water, (i) natural and semi-natural disturbed landscapes (including burned, logged and previously farmed areas), and (j) development. A map made by merging map classes into these groups is shown in **Figure 21**, and the relative abundance of map classes within each group is illustrated by **Figure 22**.

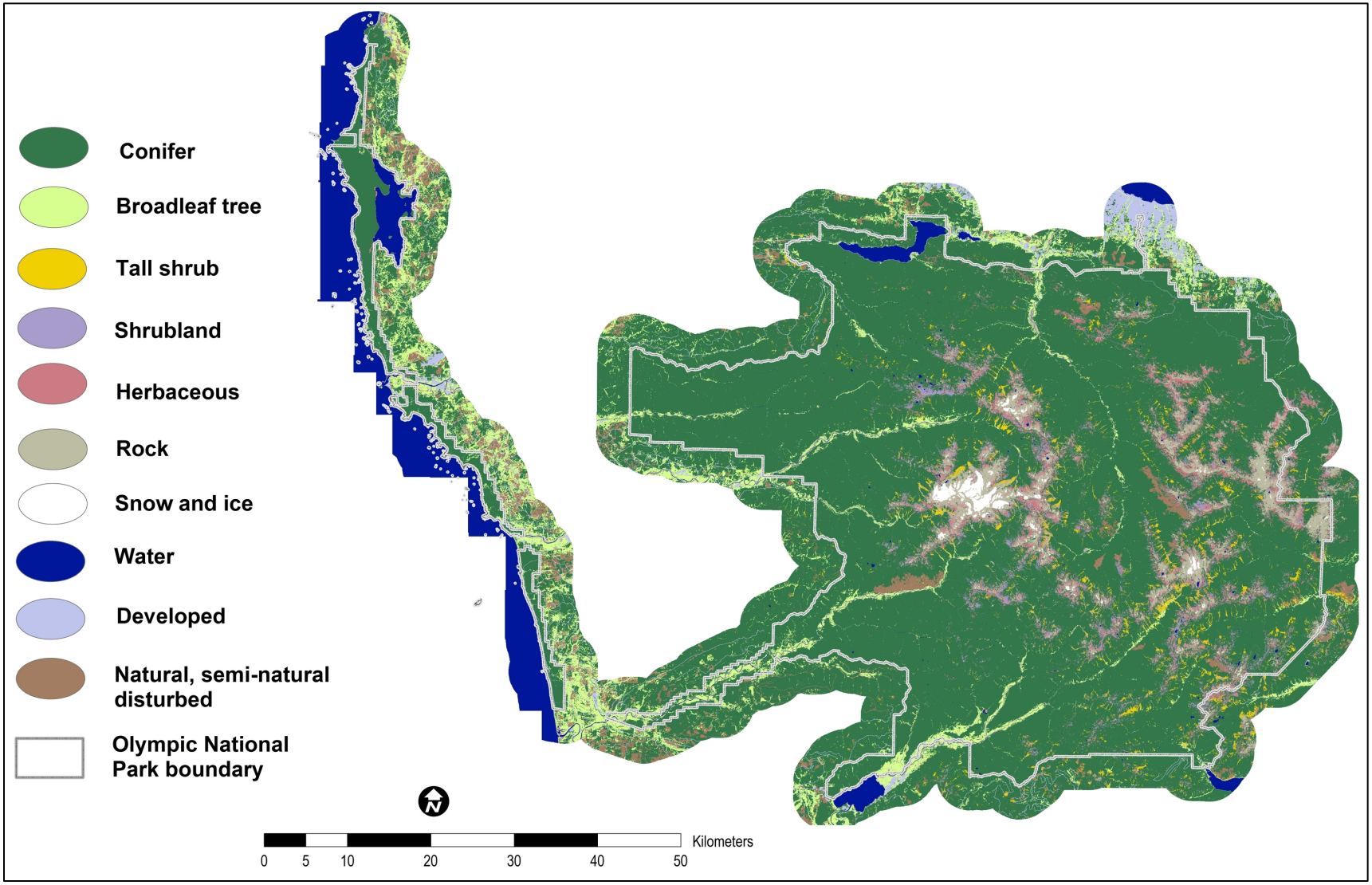


Figure 21. Lifeform map of Olympic National Park.

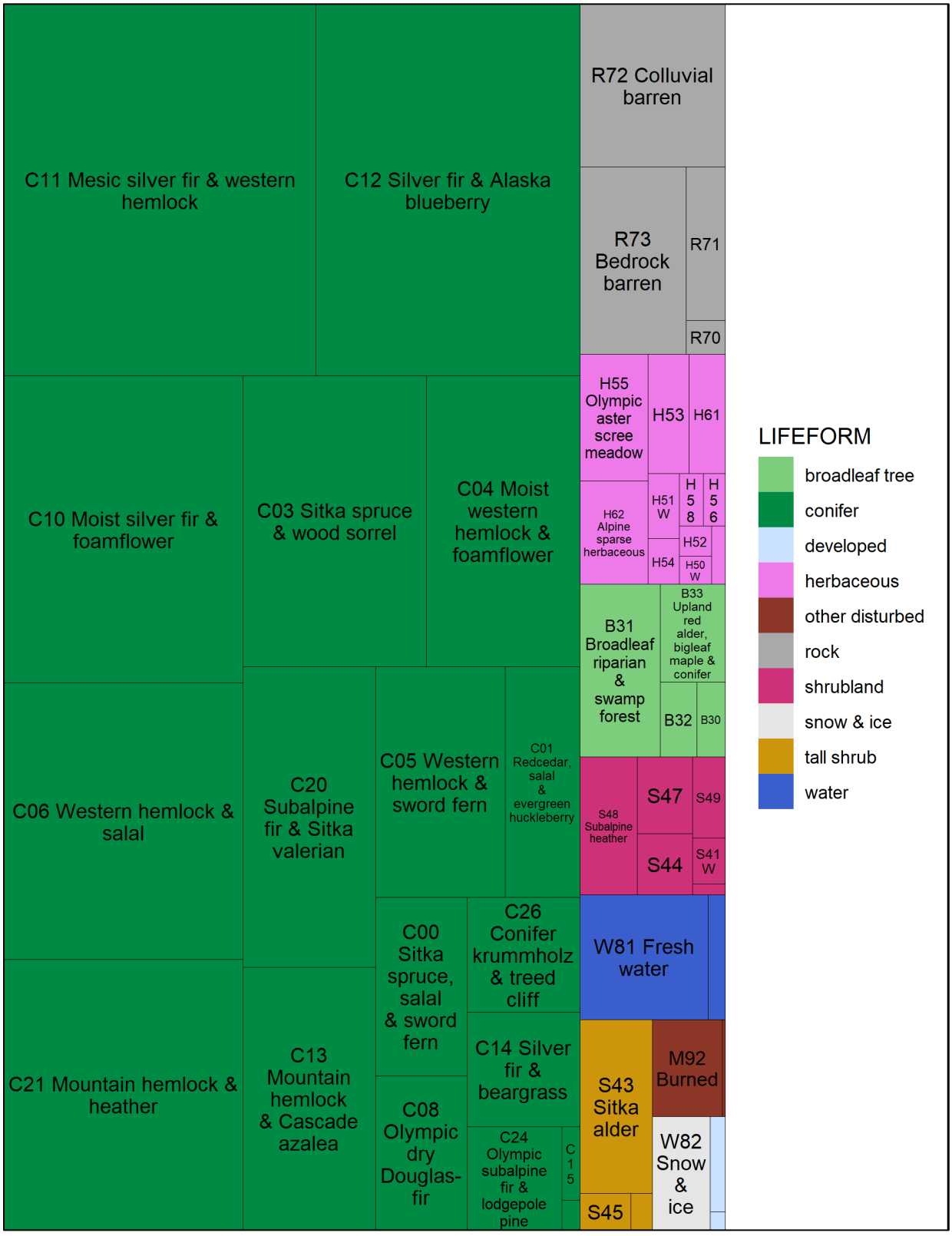


Figure 22. Relative abundance of map classes grouped by lifeform/land-use category.

Analysis of the distribution patterns of plant species provides a more fundamental (but also more detailed and complex) way of understanding vegetation patterns in the park. **Table 21** lists the most common species documented in the training plots, as well as other species that are key components of the mapped vegetation classes. Complete floristics tables are provided in INR (2021b).

Table 21. Common species in plots and other important speciesdiscussed in the text. Frequency in full-ocular training plots is given and ranked relative to all other species. Elevation zones (C=coastal, L=lowland, LM=lower montane, UM=upper montane, S=subalpine, A=alpine) with which species are most associated are marked with an 'X.' Bullets '•' indicate zones of less common but still notable occurrence. Zones where the species is not appreciably present are indicated by '–.' Scientific names follow Hitchcock and Cronquist (2018); species are listed alphabetically. INR (2021b) has a complete list.

Scientific name	Common name	Frequency	Rank	C	L	LM	UM	S	A
Abies amabilis	silver fir	26.5%	7	–	•	X	X	•	–
Abies grandis	grand fir	4.0%	98	–	X	–	–	–	–
Abies lasiocarpa	subalpine fir	22.5%	10	–	–	•	X	X	X
Acer circinatum	vine maple	13.9%	24	–	X	•	–	–	–
Acer macrophyllum	bigleaf maple	10.9%	34	–	X	–	–	–	–
Achillea millefolium	yarrow	10.1%	36	–	•	•	•	X	X
Achlys triphylla	vanillaleaf	18.5%	13	–	X	X	•	–	–
Alnus rubra	red alder	13.6%	25	X	X	–	–	–	–
Alnus viridis	Sitka alder	5.1%	76	–	•	X	X	–	–
Amelanchier alnifolia	western serviceberry	3.0%	124	–	X	X	X	–	–
Antennaria lanata	woolly pussytoes	2.1%	152	–	–	–	–	X	X
Arbutus menziesii	Pacific madrone	1.2%	224	–	X	–	–	–	–
Athyrium distentifolium	alpine lady fern	0.6%	283	–	–	–	X	–	X
Athyrium filix-femina	lady fern	15.9%	19	X	X	X	–	–	–
Berberis nervosa	dwarf Oregon-grape	16.4%	18	–	X	•	–	–	–
Bistorta bistortoides	American bistort	12.5%	28	–	–	–	–	X	X
Callitropsis nootkatensis	Alaska-cedar	17.5%	15	–	–	•	X	X	•
Carex aquatilis	water sedge	1.3%	204	•	•	–	–	X	–
Carex nigricans	black alpine sedge	3.4%	108	–	–	–	–	•	X
Carex obnupta	slough sedge	2.7%	134	X	–	–	–	–	–
Carex spectabilis	showy sedge	11.6%	33	–	–	–	–	X	X
Cassiope mertensiana	white mountain-heather	5.0%	78	–	–	–	–	•	X
Circaea alpina	enchanter's nightshade	5.8%	66	–	X	•	–	–	–
Claytonia sibirica	candyflower	8.3%	43	X	X	X	–	–	–
Clintonia uniflora	queen's cup	12.4%	29	–	•	X	X	–	–
Douglasia laevigata	cliff primrose	1.7%	175	–	–	–	–	–	X
Epilobium anagallidifolium	alpine willowherb	2.0%	157	–	–	–	X	•	X
Erythronium montanum	avalanche lily	12.5%	27	–	–	–	X	X	–

Table 21 (continued). Common species in plots and other important species discussed in the text. Frequency in full-ocular training plots is given and ranked relative to all other species. Elevation zones (C=coastal, L=lowland, LM=lower montane, UM=upper montane, S=subalpine, A=alpine) with which species are most associated are marked with an 'X.' Bullets '•' indicate zones of less common but still notable occurrence. Zones where the species is not appreciably present are indicated by '–.' Scientific names follow Hitchcock and Cronquist (2018); species are listed alphabetically. INR (2021b) has a complete list.

Scientific name	Common name	Frequency	Rank	C	L	LM	UM	S	A
Eucephalus paucicapitatus	Olympic aster	4.1%	93	–	–	–	X	X	X
Frangula purshiana	Cascara buckthorn	4.3%	90	X	•	–	–	–	–
Gaultheria shallon	salal	27.8%	6	X	•	•	–	–	–
Heracleum maximum	cow parsnip	6.5%	57	•	–	X	–	•	–
Juncus parryi	Parry's rush	3.5%	105	–	–	–	–	X	X
Juniperus communis	common juniper	6.4%	59	–	–	–	–	•	X
Kalmia microphylla	western bog-laurel	0.8%	273	X	–	–	–	•	–
Linnaea borealis	twinflower	15.8%	20	•	X	X	•	–	–
Lomatium nudicaule	Cascade desert-parsley	0.3%	389	–	X	–	–	–	•
Luetkea pectinata	partridgefoot	9.8%	38	–	–	–	•	X	X
Lupinus latifolius	subalpine lupine	16.9%	16	–	–	–	•	X	•
Luzula piperi	Piper's woodrush	1.6%	180	–	–	–	–	–	X
Lysichiton americanus	skunk-cabbage	6.0%	62	X	–	–	–	–	–
Maianthemum dilatatum	wild lily-of-the-valley	12.6%	26	X	–	–	–	–	–
Malus fusca	Pacific crabapple	1.9%	160	X	–	–	–	–	–
Oemleria cerasiformis	osoberry	0.6%	290	X	X	–	–	–	–
Orthilia secunda	one-sided wintergreen	7.6%	46	–	•	X	X	–	–
Oxalis oregana	Oregon wood-sorrel	9.4%	41	X	X	–	–	–	–
Paxistima myrsinites	Oregon-box	5.3%	72	–	X	X	X	X	•
Phacelia hastata	silver-leaf phacelia	4.0%	99	–	–	•	•	X	X
Phlox diffusa	spreading phlox	7.6%	45	–	–	–	•	X	X
Phyllodoce empetriformis	pink mountain-heather	12.4%	31	–	–	–	•	X	X
Picea sitchensis	Sitka spruce	18.0%	14	X	•	–	–	–	–
Pinus albicaulis	whitebark pine	1.0%	236	–	–	–	•	X	X
Pinus contorta	lodgepole pine	1.7%	176	–	•	X	•	–	–
Pinus monticola	western white pine	4.8%	82	–	X	X	X	–	–
Polystichum munitum	sword fern	33.9%	4	X	X	•	–	–	–
Populus trichocarpa	black cottonwood	1.6%	184	–	X	–	–	–	–
Potentilla flabellifolia	fan-leaf cinquefoil	4.7%	84	–	–	–	–	X	•
Pseudotsuga menziesii	Douglas-fir	34.0%	3	–	X	X	•	–	–
Ranunculus repens	creeping buttercup	4.4%	89	–	X	–	–	–	–
Rhododendron albiflorum	Cascade azalea	7.2%	50	–	–	–	X	•	–

Table 21 (continued). Common species in plots and other important species discussed in the text. Frequency in full-ocular training plots is given and ranked relative to all other species. Elevation zones (C=coastal, L=lowland, LM=lower montane, UM=upper montane, S=subalpine, A=alpine) with which species are most associated are marked with an 'X.' Bullets '•' indicate zones of less common but still notable occurrence. Zones where the species is not appreciably present are indicated by '–.' Scientific names follow Hitchcock and Cronquist (2018); species are listed alphabetically. INR (2021b) has a complete list.

Scientific name	Common name	Frequency	Rank	C	L	LM	UM	S	A
Rhododendron groenlandicum	bog Labrador-tea	0.9%	254	X	–	–	–	–	–
Rhododendron menziesii	false azalea	15.2%	21	X	•	•	•	–	–
Rubus nutkanus	thimbleberry	5.2%	74	–	X	X	•	–	–
Rubus pedatus	trailing raspberry	12.4%	30	–	–	X	X	–	–
Rubus spectabilis	salmonberry	25.1%	8	X	•	X	–	–	–
Salix commutata	undergreen willow	2.3%	147	–	–	–	–	X	–
Salix sitchensis	Sitka willow	0.9%	262	–	X	X	–	–	–
Sambucus racemosa	red elderberry	5.9%	63	•	X	X	–	–	–
Scirpus microcarpus	panicled bulrush	0.5%	317	–	X	–	–	–	–
Sorbus sitchensis	Sitka mountain-ash	9.0%	42	–	–	–	X	X	–
Sphagnum spp.	peat moss	1.4%	202	X	–	–	–	–	–
Spiraea splendens	mountain spirea	5.9%	64	–	–	–	X	X	–
Struthiopteris spicant	deer fern	21.9%	11	X	•	•	–	–	–
Symphoricarpos albus	common snowberry	4.1%	97	–	X	X	–	–	–
Taxus brevifolia	Pacific yew	3.5%	106	–	X	X	–	–	–
Thalictrum occidentale	western meadowrue	4.2%	91	–	–	X	•	X	–
Thuja plicata	western redcedar	23.4%	9	X	X	X	–	–	–
Tiarella trifoliata	foamflower	19.2%	12	–	X	X	X	–	–
Tolmiea menziesii	piggyback plant	7.1%	52	–	X	X	–	–	–
Toxicoscordion venenosum	meadow death camas	0.3%	369	–	X	–	–	•	–
Tsuga heterophylla	western hemlock	53.4%	1	X	X	X	•	–	–
Tsuga mertensiana	mountain hemlock	14.1%	22	–	–	–	X	X	•
Vaccinium deliciosum	Cascade blueberry	12.0%	32	–	–	–	•	X	X
Vaccinium membranaceum	big huckleberry	16.7%	17	–	•	•	X	X	–
Vaccinium ovalifolium	Alaska blueberry	29.0%	5	X	•	X	X	–	–
Vaccinium ovatum	evergreen huckleberry	7.1%	53	X	–	–	–	–	–
Vaccinium parvifolium	red huckleberry	34.5%	2	X	X	X	–	–	–
Valeriana sitchensis	Sitka valerian	14.0%	23	–	–	–	X	X	–
Xerophyllum tenax	beargrass	7.8%	44	–	–	–	X	•	–

The distribution patterns of the most important species and map classes within each of the lifeform/land-use categories are discussed below.[72] For purposes of discussion, we treat natural abiotic areas (including rock barrens, snow and ice, and water) as a single unit, and we treat wetlands as a separate unit, despite lumping them by their dominant lifeform in the lifeform map. As above, area estimates are based on the mapped area modified by the correction factor from the AA population contingency table. Species occurrence frequencies are relative to the full-ocular training plot dataset.

4.2.1. Conifers

The dry summers and relatively warm winters of the Pacific Northwest favor the development of coniferous forest as the climax lifeform where local conditions permit. Conifer-dominated vegetation is the most abundant lifeform in the park, covering four-fifths of the area, accounting for one-third of the map classes, and ranging from temperate rainforests to snow-sculpted krummholz clinging to rocky ridges. The classes are distributed fairly evenly across elevation zones: coastal and lowland forests cover 28% of the park, lower montane forests occupy 23%, upper montane forests 15% and subalpine woodlands 13%. Alpine krummholz occupies only 1.5%. The sixteen species of conifers are supplanted by other vegetation only where disturbance, snowpack, saturated soils or lack of soil development prevent their establishment and persistence.

Coastal forests

Western hemlock (*Tsuga heterophylla*) is the most abundant plant in the dataset, occurring in over half of the full-ocular training plots. It is codominant in nearly all coastal, lowland, and montane forests throughout the park, and is ubiquitous in coastal conifer forests. Salal (*Gaultheria shallon*) and deer fern (*Struthiopteris spicant*) are omnipresent in the understory of all coastal forests, and false azalea (*Rhododendron menziesii*) and wild lily-of-the-valley (*Maianthemum dilatatum*) are nearly as common. While not uncommon inland, nowhere else do these species reach the levels of abundance they achieve here.

Forests on very moist but well-drained sites near the ocean are dominated by Sitka spruce (*Picea sitchensis*), which also extends inland in lowland rainforest valleys. At the coast, it dominates C00–SITKA SPRUCE, WESTERN HEMLOCK, SALAL AND SWORD FERN FOREST, which is also characterized by very abundant sword fern (*Polystichum munitum*), red huckleberry (*Vaccinium parvifolium*) and salmonberry (*Rubus spectabilis*). These understory species are also abundant throughout most lowland and lower montane conifer forests, the latter particularly in very moist sites.

Poorly-drained coastal areas, such as the area west of Ozette Lake, host C01–REDCEDAR, WESTERN HEMLOCK, SALAL AND EVERGREEN HUCKLEBERRY FOREST, which is also found in a band just inland from C00 elsewhere. Western redcedar (*Thuja plicata*) is a widespread lowland conifer which is especially abundant in these stands. Dense thickets of evergreen huckleberry (*Vaccinium ovatum*) with skunk-cabbage (*Lysichiton americanus*) and occasional Pacific crabapple (*Malus fusca*) are

[72] The map class descriptions (Nielsen et al. 2021c) contain greater detail about the species composition, habitat and distribution associated with each map class.

characteristic in the understory here, as well as in the early-successional variant of these forests, C02–REDCEDAR, LABRADOR-TEA, SLOUGH SEDGE AND SPHAGNUM BOG.

Lowland forests

Approximately one-quarter of the park is occupied by interior lowland forests. The transition to these forests from the coastal plain occurs by way of the large west-side river valleys. Forests along the lower Queets, where the park interior most closely approaches the coast, are floristically ambiguous and difficult to distinguish from C00–SITKA SPRUCE, WESTERN HEMLOCK, SALAL AND SWORD FERN FOREST. However, a transition into a Sitka spruce-dominated valley bottom type, C03–SITKA SPRUCE, WESTERN HEMLOCK AND WOOD-SORREL FOREST, becomes evident farther upstream with a decline in coastal shrub species and the appearance of species such as Oregon wood-sorrel (*Oxalis oregana*), foamflower (*Tiarella trifoliata*), and eventually vine maple (*Acer circinatum*) and Douglas-fir (*Pseudotsuga menziesii*). The temperate rainforests in the west-side river valleys, which are often wreathed in coastal fog, are among the park's most iconic features. Douglas-fir is increasingly abundant with greater distance from the coast, and remains a canopy dominant throughout most inland forests in the park, even at subalpine elevations on the east side.

Sitka spruce vanishes outside the coastal fog zone, and all other lowland conifer forests are dominated by the pioneer species Douglas-fir and shade-tolerant western hemlock and western redcedar. Valley bottoms and well-watered toe slopes, typically in the park interior, host C04–MOIST WESTERN HEMLOCK, DOUGLAS-FIR AND FOAMFLOWER FOREST. These forests have a lush understory, with abundant foamflower, diverse ferns including sword fern and lady fern (*Athyrium filix-femina*), vanillaleaf (*Achlys triphylla*), and a variety of other moist-site species. Mesic settings have the frequently successional forest type C05–WESTERN HEMLOCK, DOUGLAS-FIR AND SWORD FERN FOREST and, on higher slope positions with coarser-textured soils, C06–WESTERN HEMLOCK, DOUGLAS-FIR AND SALAL FOREST. Western white pine (*Pinus monticola*) is a frequent associate in the latter type, but is scattered to considerably higher elevations as well. In these mesic to dry-mesic stands, dwarf Oregon-grape (*Berberis nervosa*) joins red huckleberry, salal, and often twinflower (*Linnaea borealis*) in an understory that can vary widely in density. The driest lowland forests, C08–OLYMPIC DRY DOUGLAS-FIR FOREST, are narrowly distributed in the heart of the Olympic rain shadow, particularly in the lower reaches of the Elwha River valley and around Lake Crescent. These dry stands often have prominent grand fir (*Abies grandis*) and also host the highest park densities of the evergreen broadleaf Pacific madrone (*Arbutus menziesii*) and the tall shrub western serviceberry (*Amelanchier alnifolia*), though the latter is very infrequent relative to its abundance in the North Cascades. The small understory tree Pacific yew (*Taxus brevifolia*) frequently occurs in all the above mesic to dry lowland conifer forest types.

Lower montane forests

Nearly another quarter of the park is covered in lower montane zone forests, characterized by codominance of silver fir (*Abies amabilis*) and western hemlock. Entry to this zone is marked by increased canopy prominence of silver fir, the primary successional species throughout all but the driest of mid-elevation forests. C10–MOIST SILVER FIR, WESTERN HEMLOCK AND FOAMFLOWER FOREST, the third most abundant map class in the park, occurs in valleys, on moist toe slopes and in

other protected sites. Many stands are older and contain no tree species other than silver fir and western hemlock. Alaska blueberry (*Vaccinium ovalifolium*) is the dominant understory species here. The distribution of this species is bimodal; it is also quite prominent in coastal forests but infrequent in the intervening lowlands. It is especially abundant in these wet-mesic settings where it is usually joined by foamflower, trailing raspberry (*Rubus pedatus*) and queen's cup (*Clintonia uniflora*).

West-side valley walls and other lower montane sites with more coarsely textured soils are occupied by C11–MESIC SILVER FIR AND WESTERN HEMLOCK FOREST, the most widespread forest type in the park. Stands are often successional and usually contain older and larger Douglas-firs which may form a distinct upper canopy. When silver fir and hemlock are very dense, the understory can be extremely sparse, limited to very small amounts of red huckleberry, Alaska blueberry, and an assortment of other small shade-tolerant species. Despite the extensive range of this class, it is not easily accessible due to its association with mid-slope valley walls where trails rarely go.

In the eastern half of the park, lower montane forests may include C15–LODGEPOLE PINE AND DOUGLAS-FIR WOODLAND, which is found on old terraces above the Elwha River and bedrock-limited sites on south-facing slopes on valley walls. Lodgepole pine (*Pinus contorta*) has a peculiar distribution in the park; it is most abundant in these lower montane sites, but also appears in subalpine woodlands and is scattered on coastal dunes. In the northeast corner of the park, the lower montane zone is essentially eliminated, squeezed between lowland forests and the upper montane zone.

Upper montane forests

Upper montane forests, which occupy one-eighth of the park, are transitional to the subalpine zone above. Here, the high elevation species mountain hemlock (*Tsuga mertensiana*), subalpine fir (*Abies lasiocarpa*) and Alaska-cedar (*Callitropsis nootkatensis*) become prominent in closed forests usually dominated by silver fir.

On the west side, the lush C12–SILVER FIR, HEMLOCK AND ALASKA BLUEBERRY FOREST is transitional from the lower montane, with high elevation tree species usually subordinate to silver fir, western hemlock and/or Douglas-fir. It is the second most abundant map class in the park—considerably more extensive than it is at other NCCN parks—and is marked by a mingling of silver fir with both hemlock species and Alaska-cedar, with a lush transitional understory of shrubs and forbs such as avalanche lily (*Erythronium montanum*). At higher elevations, especially on north-facing slopes, C12 gives way to the very abundant C13–MOUNTAIN HEMLOCK, SILVER FIR AND CASCADE AZALEA FOREST. Alaska-cedar is regularly prominent and occasionally codominant in the west-side upper montane, showing a preference for wetter forests, especially those regenerated after avalanche disturbance. Big huckleberry (*Vaccinium membranaceum*) replaces Alaska blueberry as the dominant understory species here, and is equally prominent on the west and east sides, extending upwards well into the subalpine zone. Cascade azalea (*Rhododendron albiflorum*) is strongly associated with this forest type and usually rivals big huckleberry for dominance in the understory.

On the east side, while cooler aspects show patterns similar to the west side, C14–SILVER FIR, BIG HUCKLEBERRY AND BEARGRASS FOREST becomes important on south-facing midslopes. This forest

type, often establishing after fires, is much more abundant in the Cascades to the southeast, such as around Mount Rainier. It is marked by codominant silver fir, western hemlock and Douglas-fir with occasional subalpine tree species and a dry-site understory that includes species such as beargrass (*Xerophyllum tenax*) and one-sided wintergreen (*Orthilia secunda*).

Subalpine forests and woodlands

Structural changes mark the passage to the subalpine zone, where forests and woodlands occupy another eighth of the park. The transition begins with closed forests becoming interrupted with shrubby openings where deep snow cover inhibits tree establishment. At higher elevations, subalpine parkland consists of forest and woodland patches alternating with increasingly large areas of open shrubland and meadow. Higher yet, snow accumulation limits the growth of conifers entirely, other than C26–CONIFER KRUMMHOLZ AND TREED CLIFF in exposed locations which are swept clear by wind.

The most abundant subalpine tree-dominated type is C21–MOUNTAIN HEMLOCK, SUBALPINE FIR AND HEATHER WOODLAND, which at the lowest elevations consists of small shrubby openings of pink mountain-heather (*Phyllodoce empetriformis*) and Cascade blueberry (*Vaccinium deliciosum*) in a matrix of mountain hemlock. At higher elevations and in more exposed areas, the trees become limited to clumps known as tree islands. The type is better developed in the moister west-side mountains, but is found on the east side as well, where subalpine fir becomes an increasingly important component.

C20–SUBALPINE FIR AND SITKA VALERIAN FOREST AND WOODLAND is transitional from the upper montane zone and is found throughout the central and eastern portions of the park. These forests are dominated by subalpine fir, often with silver fir and Alaska-cedar, and usually with a lush understory and a diverse herbaceous component including Sitka valerian (*Valeriana sitchensis*). Moister occurrences of this type, especially forests developed following avalanche disturbance, may have dominant Alaska-cedar. The type is particularly abundant in the northeast park section on north- and west-facing slopes. Subalpine fir is shade-intolerant and a rapid colonizer of suitable habitat, and it extends to lower elevations in this area, where abrupt transitions may occur from dry lowland forests. C24–OLYMPIC SUBALPINE FIR AND LODGEPOLE PINE WOODLAND is prominent on higher elevation south-facing slopes on the east side. These dry open woodlands often include lodgepole pine and Douglas-fir. Whitebark pine (*Pinus albicaulis*) may occur here, but is rare in OLYM compared to the other NCCN parks.

4.2.2. Broadleaf trees

Communities dominated by deciduous broadleaf trees occupy three percent of the park and are especially prominent on the valley bottom floodplains of west-side river systems. They also are present on slopes recovering from landslides, fire or logging.

Floodplains

The colonization phase of floodplain successional dynamics is represented by B30–SUCCESSIONAL GRAVEL BAR SHRUBLAND. The dominant woody plants in these communities are red alder (*Alnus rubra*), black cottonwood (*Populus trichocarpa*), and Sitka willow (*Salix sitchensis*). Without

repeated disturbance, these shrublands mature into B31–BROADLEAF RIPARIAN AND SWAMP FOREST. These forests occur mostly in the broad west-side valleys and cover 1.6% of the park, double their abundance in the other NCCN parks. Like C03–SITKA SPRUCE, WESTERN HEMLOCK AND WOOD-SORREL FOREST, which often occurs nearby, salmonberry and sword fern are very abundant in these forests. Unlike the Sitka spruce forest, piggyback plant (*Tolmiea menziesii*) and candyflower (*Claytonia sibirica*) are very common as well. The understory is often heavily impacted by elk grazing; the introduced plant creeping buttercup (*Ranunculus repens*) is common as well. These forests are also the primary habitat in the park for the tall shrubs Cascara buckthorn (*Frangula purshiana*) and osoberry (*Oemleria cerasiformis*), though they are not abundant.

Uplands

Two other deciduous forest types occur in the park. B32–COASTAL RED ALDER FOREST is a successional type that emerges between Sitka spruce forests following stand-replacing landslide disturbance on coastal bluffs. Its understory is similar to that of the undisturbed forests adjacent, but accompanied by additional opportunists such as cow parsnip (*Heracleum maximum*). B33–UPLAND RED ALDER, BIGLEAF MAPLE AND CONIFER FOREST is characteristic of disturbed areas in the lowland conifer zone, occurring on toe-slope debris aprons and in stands regenerated after logging. Bigleaf maple (*Acer macrophyllum*) is the dominant broadleaf tree in these upland forests, but red alder is usually present as well. Again, the understory is usually similar to adjacent conifer forests, but in this case is supplemented by delicate forbs such as enchanter's nightshade (*Circaea alpina*).

4.2.3. Upland tall shrubs

Tall shrublands in uplands cover about two percent of the park, occupying avalanche tracks, montane talus slopes, riparian benches and toe-slope debris aprons. Avalanches tend to impact the same slopes year after year, carving out chutes through montane and subalpine forests. The regular disturbances favor resilient and rapidly resprouting shrubs rather than tall and brittle trees. Sitka alder (*Alnus viridis*) bends rather than breaks when walloped by snow, and thrives in avalanche zones and moist talus. Although the species is not extremely common in the dataset, it is generally dominant where it occurs, and S43–SITKA ALDER SHRUBLAND is the most abundant non-forest vegetation type in the park. Salmonberry, red elderberry and lady fern are common associates.

Other tall shrubland classes occur on avalanche toe slopes, where Sitka alder gives way to a variety of other shrub species. Often found just above large streams and rivers, S42–SITKA WILLOW RIPARIAN SHRUBLAND forms dense thickets that are codominated by Sitka willow, Sitka alder and salmonberry with a mix of moisture-loving forbs from lowlands and middle elevations. Drier talus and avalanche toe slopes host S45–VINE MAPLE SHRUBLAND, in which vine maple (*Acer circinatum*) is often joined by young bigleaf maple and a variety of smaller shrubs and herbaceous plants. These sites have strong floristic similarities to B33–UPLAND RED ALDER, BIGLEAF MAPLE AND CONIFER FOREST. Vine maple is also very common in the understory of lowland conifer forests.

4.2.4. Upland shrublands

Shorter shrublands in uplands also cover about two percent of the park, exploiting disturbance in the montane zone and exposed areas with poor soil development at higher elevations. The lowest elevation shrublands are on lower montane toe slopes with thimbleberry (*Rubus nutkanus*) or

common snowberry (*Symphoricarpos albus*) and an assortment of tall forbs; these are mapped as S44–THIMBLEBERRY SHRUBLAND, TALL FORBS AND BRACKEN FERN. At higher elevations, areas dominated by a mix of big huckleberry, Cascade blueberry, mountain spirea (*Spiraea splendens*) and Sitka mountain-ash (*Sorbus sitchensis*) are scattered through the upper montane and subalpine zones. These successional shrublands, mapped as S47–SUCCESSIONAL HUCKLEBERRY SHRUBLAND, typically occur between forests and subalpine meadows, and may be associated with recovery from fire.

The iconic mountain-heather shrublands are among the most abundant features of the subalpine and alpine zones in all but the northeastern quarter of the park. These dwarf-shrublands are dominated by pink mountain-heather, white mountain-heather (*Cassiope mertensiana*) and Cascade blueberry, and are particularly prominent on upper slopes with cold aspects. S48–SUBALPINE HEATHER SHRUBLAND, which usually includes prominent avalanche lily and other forbs, is typically found just above C21–MOUNTAIN HEMLOCK, SUBALPINE FIR AND HEATHER WOODLAND, but there is a substantial transition zone in which the two classes mosaic at a scale of 10–20 meters, resulting in an extensive landscape of tree islands and heather openings. Succession toward mountain hemlock is usually evident in these areas. A gradual transition to S49–ALPINE HEATHER SHRUBLAND is evident with increasing elevation and exposure. This class is characterized by more compact vegetation with considerably fewer forbs, increased abundance of white mountain-heather, and reduced abundance of Cascade blueberry.

4.2.5. Upland herbaceous vegetation

Lingering snowpack, low temperatures, desiccating winds and repeated disturbance create conditions under which only herbaceous plants can survive. A diverse assortment of upland herbaceous plant communities—including lush forb meadows, rocky graminoid meadows and sparse alpine cushion plants—share four percent of the park's area. Herbaceous communities often transition across short distances, responding to finer-scale changes in topography, substrate and soil moisture than adjacent forests.

Lowland and montane elevations

The lowest elevation herbaceous communities in the park occur in forest openings on bedrock-limited soils and are mapped as H58–BEDROCK BALDS AND SPARSELY VEGETATED FOREST OPENINGS. Balds are concentrated around the northern reaches of the Elwha River and above the Skokomish River and reflect the glacial scouring these watersheds experienced. Mosses, lichens and spikemosses are most abundant at these sites, but they also contain species such as barestem biscuitroot (*Lomatium nudicaule*) and meadow death camas (*Toxicoscordion venenosum*), which are confined to balds in the park, as well as additional herbaceous species more common in the alpine zone. Openings in lower montane forests may also occur on toe slopes, where herbaceous communities with cow parsnip, western meadowrue (*Thalictrum occidentale*) and other tall forbs persist due to habitual disturbance from landslides and avalanche runouts. These meadows, H52–COW PARSNIP MEADOW, usually occur in a mosaic with S44–THIMBLEBERRY SHRUBLAND, TALL FORBS AND BRACKEN FERN, discussed above.

Subalpine and alpine elevations

Lush meadows of showy sedge (*Carex spectabilis*), subalpine lupine (*Lupinus latifolius*), American bistort (*Bistorta bistortoides*) and Sitka valerian (*Valeriana sitchensis*) are mapped as H53–SHOWY

SEDGE & SITKA VALERIAN MEADOW. Found throughout the park, they are more abundant on the moist west side, where they are often mosaicked with mountain-heather shrublands. In rocky areas, this community may grade into H54–MOIST TALUS VEGETATION, which occurs as patches of alpine lady fern (*Athyrium distentifolium*), alpine willowherb (*Epilobium anagallidifolium*) and many other forbs within protected and often sub-irrigated portions of talus slopes. The most abundant herbaceous type in the park, H55–OLYMPIC ASTER SCREE MEADOW, occurs on more densely vegetated stabilized scree slopes, often on south-facing colluvial fans. This is a mixed vegetation characterized by dry site species such as Olympic aster (*Eucephalus paucicapitatus*) and silver-leaf phacelia (*Phacelia hastata*) but with more protected pockets hosting species typical of lusher meadows. In exposed alpine areas, cushion plants such as Tolmie's saxifrage (*Micranthes tolmiei*) and cliff primrose (*Douglasia laevigata*) join partridgefoot (*Luetkea pectinata*), Piper's woodrush (*Luzula piperi*) and other stoic plants to form H62–ALPINE SPARSE HERBACEOUS VEGETATION. This class is found throughout alpine areas on rocky upper slopes where snow lingers late into the growing season, but is most abundant in the Bailey Range and surrounding areas.

Two additional high elevation herbaceous types replace lusher meadows and mountain-heather in the dry northeastern park quadrant. H61–SPREADING PHLOX AND COMMON JUNIPER MEADOW, characterized by spreading phlox (*Phlox diffusa*), yarrow (*Achillea millefolium*) and common juniper (*Juniperus communis*), is found on south-facing subalpine ridges. It has strong similarities to H64–ALPINE LUPINE PUMICE VEGETATION at MORA, and the two classes should perhaps be combined. Common juniper is also abundant in C26–CONIFER KRUMMHOLZ AND TREED CLIFF, unsurprisingly, as krummholz is its natural growth form. Flats with coarse-textured soils that are wet in spring but very dry by late summer contain a vegetation type that is more common at Mount Rainier, H56–SUBALPINE SUMMER-DRY GRASS-FORB MEADOW. Parry's rush (*Juncus parryi*) and woolly pussytoes (*Antennaria lanata*) best distinguish this meadow type from the others.

4.2.6. Wetlands

Less than one percent of the park is mapped in one of the six wetland map classes. Low-gradient areas conducive to wetland formation are mostly limited in the park to coastal areas, lowland valleys and high-elevation headwaters basins, often where glaciers have carved out flat-bottomed cirques. There are riparian communities and occasional mid-slope seeps in the intermediate montane zones, but few large wetlands.

<u>Lowlands</u>

Some areas of the coastal strip, particularly near Lake Ozette, host shrubby wetlands dominated by bog Labrador-tea (*Rhododendron groenlandicum*), western bog-laurel (*Kalmia microphylla*), slough sedge (*Carex obnupta*), and sphagnum moss (*Sphagnum* spp.). These contain varying amounts of western redcedar and western hemlock and appear to be successional to C01–REDCEDAR, WESTERN HEMLOCK, SALAL AND EVERGREEN HUCKLEBERRY FOREST. Some occurrences appear related to past human settlement. Inland low elevation herbaceous wetlands are represented by H50W–LOW ELEVATION HERBACEOUS WETLAND; these are usually dominated by graminoids such as panicled bulrush (*Scirpus microcarpus*) and water sedge (*Carex aquatilis*), and often occur along pond and lake margins. In river backwaters and areas of beaver activity, they form dynamic complexes with

S40W–LOW ELEVATION SHRUB SWAMP wetlands dominated by Sitka willow and rose spirea (*Spiraea douglasii*), and with swampy B31–BROADLEAF RIPARIAN AND SWAMP FOREST. Herbaceous wetlands around ponds and lakes may be ringed by a shrub wetland and limited to a narrow strip between that and the water.

Subalpine and alpine

Herbaceous subalpine wetlands, mapped as H51W–SUBALPINE HERBACEOUS WETLAND, usually have prominent fan-leaf cinquefoil (*Potentilla flabellifolia*) and black alpine sedge (*Carex nigricans*) with a range of other forbs and sedges. These wetlands are found throughout the high elevations of the park on cirque floors and in other headwaters basins. Wetlands in similar environments but dominated by dense patches of undergreen willow (*Salix commutata*) are especially abundant in the rain shadow and are mapped as S41W–SUBALPINE WILLOW WETLAND. The two types often form mosaics, with undergreen willow wetlands adjacent to low-gradient streams and surrounded by herbaceous wetlands. They are usually in low spots within subalpine parkland but can also form on sloping seeps. A final high elevation wetland type consists of a dense turf of black alpine sedge and other sedges, usually with prominent partridgefoot, which occur throughout the alpine zone in depressions holding snow beds or collecting melt from above. These perched wetlands are mapped as H60W–BLACK ALPINE SEDGE WETLAND; patches are often inclusions within subalpine or alpine heather shrublands, with rock barrens usually nearby.

4.2.7. Natural abiotic areas

Unvegetated natural areas cover eight percent of the park, a much smaller proportion than at MORA and NOCA. Most of the unvegetated area is barren rock, with roughly equal amounts of R72–COLLUVIAL BARREN and R73–BEDROCK BARREN, substantially less R71–ALLUVIAL BARREN AND DEBRIS-COVERED ICE, and a small amount of R70–COASTAL BARREN. Although most of these barrens may entirely lack vascular plants, many include sparse vegetation. The most common vascular plant species in barrens as a group are partridgefoot, subalpine fir, Tolmie's saxifrage and alpine willowherb. About two percent of the park is open fresh water; much of this is found in Lake Crescent and Ozette Lake.

4.2.8. Natural and semi-natural disturbed landscapes

Nearly one percent of the park is mapped as M92–BURNED WITH UNCERTAIN VEGETATION. These are areas that have experienced severe fire in the last 35 years but do not resemble any of the fire-adapted map classes discussed above. They were not well-sampled in our fieldwork; vegetation cover is likely low and probably consists of a variety of early successional plants, including fireweed (*Chamaenerion angustifolium*).

4.2.9. Development

Roads and developed areas cumulatively occupy 0.2% of the park.

4.3. Influence of disturbance

The USNVC and the mapping classification are best developed for stable climax and late seral vegetation types, but disturbances are a major driver of vegetation composition in the park. Post-disturbance trajectories may follow consistent patterns of vegetation colonization or recovery

represented in the classification, but also may result in unique combinations of species that do not fit the classification well. Both scenarios are discussed below.

Windthrow drives significant vegetation dynamics in the lowland coniferous forests (Copass et al. 2016), altering successional pathways depending on disturbance intensity; these have not been well characterized or studied. Our fieldwork showed that, despite dramatic canopy structure changes brought about by disturbance, most windthrow patches still floristically align with undisturbed map classes because the floristics of the canopy trees and associated regeneration as well as the understory species that take over the new canopy gaps, were largely present in the species assemblage present before the disturbance. These recovering wind-thrown patches make up a minor portion of lower elevation forests.

Avalanches represent another important agent of change. They periodically shatter tree trunks, favoring shorter and more flexible tall shrub plant communities along their established paths. These communities, which repeat regularly across the upper and lower montane zones, are mapped as S43–SITKA ALDER SHRUBLAND and S45–VINE MAPLE SHRUBLAND. The less severe impacts in adjacent areas result in less consistent outcomes that are difficult to predict or map. The outer flanks of avalanche tracks often contain battered conifer forests with variable understories, and concentrated debris deposition zones in avalanche runouts often feature a haphazard and opportunistic mix of subalpine plants displaced from above with montane plants from nearby. Species composition in these areas varies from site to site and would not be easily placed in any classification.

Flooding is common along the wild rivers and streams in the park. Rivers in the lower valleys regularly change course, washing away established forests and over time resulting in a patchwork of even-aged broadleaf forests and older conifer-dominated forests. Less destructive flooding can kill standing trees or bury the understory in cobbles, leading to atypical plant communities that are not captured in the classification. In the absence of continued flooding, abandoned channels, banks, and bars are eventually colonized and follow trajectories toward climax conifer forests. Colonizing woody plants and their herbaceous associates often show consistent floristics, which we map as B30–SUCCESSIONAL GRAVEL BAR SHRUBLAND, but a wide variety of species are possible depending on propagule availability, substrate and water table depth.

Glaciers in the park have receded significantly, mostly over the Holocene (Osborn et al. 2012, Beget 1984), but also more recently (Pelto 2017). They have left lakes in their wakes, scoured bedrock benches, and piled unsorted glacial till in moraines of various ages. The vegetation (or lack thereof) on these landforms depends on age, climate, water table and propagule availability. Older moraines host subalpine plant communities such as S48–SUBALPINE HEATHER SHRUBLAND or even conifer forests. Younger moraines are less likely to host a cohesive plant community and are instead dominated by a smattering of whatever nearby plant species happen to get a toehold. These early seral assemblages are often unlike any map classes, but are mapped as their best match: for alpine moraines, this may be S49–ALPINE HEATHER SHRUBLAND or H62–ALPINE SPARSE HERBACEOUS VEGETATION; on terminal moraines in valleys it is often B30–SUCCESSIONAL GRAVEL BAR SHRUBLAND or S43–SITKA ALDER SHRUBLAND.

We created several map classes to account for land cover types not treated in the associations: areas significantly disturbed by fire or logging, non-native vegetation resulting from past agriculture, and developed and agricultural land within the park and adjacent mapped areas. These land cover types were not inventoried either due to access issues or because they do not contain native vegetation or represent significant conservation value. Although they cannot be described floristically, they were mapped to general land-cover/land-use categories to prevent gaps in the map coverage.

The most abundant of these types within the park are recently burned areas. Sites burned within the last 35 years that do not resemble typical post-fire vegetation classes were mapped as M92–BURNED WITH UNCERTAIN VEGETATION. Variability in pre-fire vegetation, fire severity, and propagule availability result in diverse recovery pathways, limiting our capability to map these areas with more specificity.

Olympic National Park's origin as a federal forest reserve coupled with its relative inaccessibility mostly protected forests from clearcutting prior to park establishment. Nonetheless, a few places within the current park boundary saw timber harvest, most notably in the valleys of the Elwha and Sol Duc Rivers. The Elwha floodplain at the former Lake Mills reservoir was significantly logged prior to damming (Crane 2011). Regenerated cuts within the park are handled well by the association-level classification. More recently logged areas outside the park boundary that could not be confidently assigned to one of the other map classes were coded as M93–TIMBERLAND WITH UNCERTAIN VEGETATION. Atypical conditions are found in these areas due to forest management practices such as replanting and the use of herbicides to suppress growth of deciduous trees and shrubs (Washington DNR 2018).

Formerly farmed areas were mapped as M90–RUDERAL MEADOW. These areas are likely to continue to slowly recover into semi-natural herbaceous meadows and shrublands or forest as trees encroach (Acker et al. 2014). M94–DEVELOPMENT includes developed sites within the park such as visitor centers, housing, and maintenance facilities. It also includes several nearby communities as well as various farmed, residential, and industrial lands outside the park. Park roads are mapped as M95–ROADS IN PARK.

4.4. Guidelines for map use

Before using the map products, users should thoroughly review both the map class descriptions and the accuracy assessment. The map represents existing vegetation as of summer 2015, although it may reflect the impacts of disturbance occurring before August 4, 2017. Vegetation patches smaller than 500 m^2 may not appear in the map; patches smaller than 90 m^2 are definitely not captured. Narrow ribbon-like artifacts may be present near transitions between distinct lifeforms. In order to capture real vegetation that occurs in elongated slender patches, we did not aggressively filter these artifacts.

For some map uses, the fine floristic distinctions between our map classes will likely be unnecessary. We've provided some guidance on merging map classes into dominant lifeform groups, but likely other combinations will be useful. When combining map classes into broader categories (e.g. silver fir forests, mountain-heather dwarf-shrublands), consider floristic similarity, spatial proximity (e.g.,

“are the classes typically found adjacent on the ground?”), and confusion (e.g., “how confused are the classes in the accuracy assessment?”).

Planning of management or monitoring activities based on the vegetation map should always incorporate a consideration of the assessed accuracy of the map classes involved. Whether user’s or producer’s accuracy is a more appropriate metric depends on the issue. If a monitoring study requires field sampling within a map class, the class user’s accuracy should be considered before sending crews to randomly generated locations. For map classes with lower user’s accuracies, additional steps should be taken—at a minimum, examining recent aerial imagery—to ensure the sample locations are indeed occupied by the target class. On the other hand, the practicality of delineating the spatial bounds of a vegetation type is a function of the class producer’s accuracy. Map classes with low producer’s accuracy are not mapped in many places where they are present, so their distribution will be less clear. In some cases, an application might require consideration of the full population contingency table. For instance, an assessment of the impacts of a mapped disturbance event on habitat availability would need to estimate the fractional composition of map classes in the disturbed area. Although a simple summary based on the mapped classes would be easy, a better approach might be to apply area estimate corrections based on the population contingency table, as was done to estimate the map class areas for **Table 20**.

Literature Cited

Acker, S. A., M. D. Tetreau, and D. W. Allen. 2014. Invasive plants in the Queets Valley, Olympic National Park: Former homesteads and surrounding watershed. Natural Resource Technical Report NPS/OLYM/NRTR—2014/898. National Park Service, Fort Collins, Colorado.

Agee, J. K. 1993. Fire ecology of Pacific Northwest forests. Island Press, Washington, District of Columbia.

Anderson, K. H. 2009. The Ozette prairies of Olympic National Park: their former indigenous uses and management. Report to Olympic National Park, Port Angeles, Washington.

Andrews, H. J., and R. W. Cowlin. 1936. Forest type map, state of Washington, southeast quarter. USDA Forest Service, Pacific Northwest Forest Experiment Station, Forest Survey Staff. Portland, Oregon.

Baccus, W. D. 2018. A park interpreter's guide to the climate of Hurricane Ridge, Olympic National Park: climate summary for water years 2000 to 2017. Natural Resource Report NPS/NCCN/NRR— 2018/1714. National Park Service, Fort Collins, Colorado.

Barnes, E., T. Clarke, S. Richards, P. Colaizzi, J. Haberland, M. Kostrzewski, P. Waller, C. Choi, E. Riley, and T. Thompson. 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. Pages 16–19 *in* University of Minnesota. Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, Minnesota.

Beget, J. E. 1984. Tephrochronology of Late Wisconsin deglaciation and Holocene Glacier fluctuations near Glacier Peak, North Cascade Range, Washington. Quaternary Research 21:304–316.

Bishop, C. M. 2006. Pattern recognition and machine learning. Springer, New York.

Bivand, R., T. Keitt, and B. Rowlingson. 2014. R package 'rgdal' version 0.8-16. Available at: https://cran.r-project.org/web/packages/rgdal/index.html (accessed 01 July 2020).

Breiman, L. 2001. Random forests. Machine Learning 45:5–32.

Brunner, R. L., E. M. Nielsen, and C. Copass. 2017. A data-driven method for assembling map classes from vegetation associations in Washington National Parks. Organized poster session, Ecological Society of America Conference, Portland, Oregon. Abstract available at: https://eco.confex.com/eco/2017/webprogram/Paper62429.html.

Brunner, R. L., C. Copass, L. K. Wise, and E. M. Nielsen. 2021. Olympic National Park map class key. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Buckingham, N. M., E. G. Schreiner, T. N. Kaye, J. E. Burger and E. L. Tisch. 1995. Flora of the Olympic peninsula. Northwest Interpretive Association and Washington Native Plant Society, Seattle, Washington.

Burke Herbarium. 2020. Burke herbarium image collection. Available at: https://biology.burke.washington.edu/herbarium/imagecollection.php (accessed 30 December 2020).

Cao, Q., Y. Miao, H. Wang, S. Huang, S. Cheng, R. Khosla, and R. Jiang. 2013. Non-destructive estimation of rice plant nitrogen status with Crop Circle multispectral active canopy sensor. Field Crops Research 154:133–144.

Chander, G., B. L. Markham, and D. L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+ and EO-1 ALI sensors. Remote Sensing of Environment 113:893–903.

Chavez, P. S. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sensing of Environment 24:459–479.

Cibula, W. G. 1981. Computer implemented land cover classification using Landsat MSS digital data: a cooperative research project between the National Park Service and NASA. Report 193. NASA Earth Resources Laboratory, Bay Saint Louis, Mississippi.

Cibula, W. G., and M. O. Nyquist. 1987. Use of topographic and climatological models in a geographical data base to improve Landsat MSS classification for Olympic National Park. Photogrammetric Engineering and Remote Sensing 53:67–75.

Cohen, W. B., and T. A. Spies. 1992. Estimating structural attributes of Douglas-fir–western hemlock forest stands from Landsat and SPOT imagery. Remote Sensing of Environment 41:1–17.

Cohen, W. B., T. K. Maiersperger, S. T. Gower, and D. P. Turner. 2003. An improved strategy for regression of biophysical variables and Landsat ETM+ data. Remote Sensing of Environment 84:561–571.

Congalton, R. G., and K. Green. 1999. Assessing the accuracy of remotely sensed data: principles and practices. Lewis Publishers, Boca Raton, Louisiana.

Conrad, O., B. Bechtel, M. Bock, H. Dietrich, E. Fischer, L. Gerlitz, J. Wehberg, V. Wichmann, and J. Böhner. 2015. System for Automated Geoscientific Analyses (SAGA) version 2.1.4. Geoscientific Model Development 8:1991–2007.

Copass, C., and T. Ramm-Granberg. 2016a. Ebey's Landing National Historic Reserve vegetation inventory and mapping project. Natural Resource Report NPS/NCCN/NRR—2016/1127. National Park Service, Fort Collins, Colorado.

Copass, C., and T. Ramm-Granberg. 2016b. Vancouver National Historic Reserve vegetation inventory and mapping project. Natural Resource Report NPS/NCCN/NRR—2016/1128. National Park Service, Fort Collins, Colorado.

Copass, C., N. Antonova, and S. Clary. 2016. Landsat-based monitoring of landscape dynamics in Olympic National Park 1985–2010. Natural Resource Data Series. NPS/NCCN/NRDS 2016/1053. National Park Service, Fort Collins, Colorado.

Crane, J. 2011. Finding the river: and environmental history of the Elwha. Oregon State University Press. Corvallis, Oregon.

Crawford, R. C., C. B. Chappell, C. C. Thompson, and F. J. Rocchio. 2009. Vegetation classification of Mount Rainier, North Cascades and Olympic National Parks. Natural Resource Technical Report NPS/NCCN/NRTR—2009/D-586. National Park Service. Fort Collins, Colorado. Available at: https://irma.nps.gov/DataStore/Reference/Profile/661669 (accessed 01 July 2020).

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology 88:2783–2792.

Daughtry, C., C. Walthall, M. Kim, E. B. De Colstoun, and J. McMurtrey. 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment 74:229–239.

Davey, C. A., K. T. Redmond, and D. B. Simeral. 2006. Weather and Climate Inventory, National Park Service, North Coast and Cascades Network. Natural Resource Technical Report NPS/ NCCN/NRTR—2006/010. National Park Service, Fort Collins, Colorado.

De Cáceres, M., and P. Legendre. 2009. Associations between species and groups of sites: indices and statistical inference. Ecology 90:3566–3574. Code available at: https://cran.r-project.org/web/packages/indicspecies/index.html (accessed 01 July 2020).

De Cáceres, M., X. Font, and F. Oliva. 2010. The management of vegetation classifications with fuzzy clustering. Journal of Vegetation Science 21:1138–1151. Code available at: https://cran.r-project.org/web/packages/vegclust/index.html (accessed 01 July 2020).

Dodwell, A. D., and T. F. Rixon. 1902. Forest conditions in the Olympic Forest Reserve, Washington, from notes by Arthur Dodwell and Theodore F. Rixon. U. S. Geological Survey Professional Paper. Government Printing Office, Washington, District of Columbia.

Esri. 2013. ArcGIS Desktop 10.2. Environmental Systems Research Institute, Redlands, California.

Evans, J. S., and S. A. Cushman. 2009. Gradient modeling of conifer species using random forests. Landscape Ecology 24:673–683.

Federal Geographic Data Committee (FGDC). 2008. National vegetation classification standard, Version 2, FGDC-STD-005-2008. Available at: https://www.fgdc.gov/standards/projects/FGDC-standards-projects/vegetation/NVCS_V2_FINAL_2008-02.pdf (accessed 01 July 2020).

Fonda, R. W., and L. C. Bliss. 1969. Forest vegetation of montane and subalpine zones, Olympic Mountains, Washington. Ecological Monographs 39:271–301.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment 80:185–201.

Franklin, J. F., and C. T. Dyrness. 1988. Natural vegetation of Oregon and Washington. Oregon State University Press, Corvallis, Oregon.

Franklin, J. F., W. H. Moir, M. A. Hemstrom, S. E. Greene, and B. G. Smith. 1988. The forest communities of Mount Rainier National Park. Scientific Monograph Series No. 19. U. S. Department of Interior, National Park Service, Washington, District of Columbia.

Grossman, D. H., D. Faber-Langendoen, A. S. Weakley, M. Anderson, P. Bourgeron, R. Crawford, K. Goodin, S. Landaal, K. Metzler, K. D. Patterson, M. Pyne, M. Reid, and L. Sneddon. 1998. International classification of ecological communities: terrestrial vegetation of the United States. Volume I. The National Vegetation Classification System: development, status, and applications. The Nature Conservancy, Arlington, Virginia.

Hall, D. K., G. A. Riggs, and V. V. Salomonson. 1995. Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. Remote Sensing of Environment 54:127–140.

Healey, S. P., W. B. Cohen, Z. Yang, and O. N. Krankina. 2005. Comparison of tasseled cap-based Landsat data structures for use in forest disturbance detection. Remote Sensing of Environment 97:301–310.

Henderson, J. A., R. D. Lesher, D. H. Peter, and C.D. Ringo. 2011. A landscape model for predicting potential natural vegetation of the Olympic Peninsula USA using boundary equations and newly developed environmental variables. General Technical Report PNW-GTR-841. USDA Forest Service, Pacific Northwest Research Station, Portland, Oregon.

Henderson, J. A., E. H. Peter, R. D. Lesher, and D. C. Shaw. 1989. Forest plant associations of the Olympic National Forest. R6 Ecological Technical Paper 001-88. USDA Forest Service Pacific Northwest Region, Portland, Oregon.

Hijmans, R. J. 2018. R package 'raster': geographic data analysis and modeling, version 2.8-4. Available at: https://cran.r-project.org/web/packages/raster/index.html (accessed 01 July 2020).

Hitchcock, C. L., and A. Cronquist. 1973. Flora of the Pacific Northwest, first edition. University of Washington Press, Seattle, Washington.

Hitchcock, C. L., and A. Cronquist. 2018. Flora of the Pacific Northwest, second edition. Burke Museum Herbarium & University of Washington Press, Seattle, Washington.

Houston, D. B., E. G. Schrenier, and B. B. Moorhead. 1994. Mountain goats in Olympic National Park: biology and management of an introduced species. National Park Service Scientific Monograph Series. NPS/NROLYM/NRSM-94/25. National Park Service, Denver, Colorado.

Houze, R. A., L. A. McMurdie, W. A. Petersen, M. R. Schwaller, W. Baccus, J. D. Lundquist, C. F. Mass, B. Nijssen, S. A. Rutledge, D. R. Hudak, S. Tanelli, G. G. Mace, M. R. Poellot, D. P. Lettenmaier, J. P. Zagrodnik, A. K. Rowe, J. C. Dehart, L. E. Madaus, and H. C. Barnes. 2017. The Olympic Mountains experiment (OLYMPEX). American Meteorological Society 98:2167–2188.

Huang, C., B. Wylie, L. Yang, C. Homer, and G. Zylstra. 2002. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. International Journal of Remote Sensing 23:1741–1748.

Institute for Natural Resources (INR). 2021a. Olympic National Park accuracy assessment contingency tables. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Institute for Natural Resources (INR). 2021b. Taxonomic tables for Mount Rainier, Olympic and North Cascades National Parks. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2283943.

Isaaks, E. H., and R. M. Srivastava. 1988. Spatial continuity measures for probabilistic and deterministic geostatistics. Mathematical Geology 20:313–341.

Jones, G. N. 1947. A Botanical Survey of the Olympic Peninsula, Washington. University of Washington, Seattle, Washington.

Kagan, J. S., E. M. Nielsen, M. D. Noone, J. C. van Warmerdam, L. K. Wise, G. Kittel, and C. Copass. 2012. Lewis and Clark National Historic Park vegetation classification and mapping project report. Natural Resource Report NPS/LEWI/NRR—2012/597. National Park Service, Fort Collins, Colorado.

Kauth, R. J., and G. S. Thomas. 1976. The tasseled cap: a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data: 4B41–4B51. Purdue University, West Lafayette, Indiana.

Key, C. H., and N. C. Benson. 2002. Measuring and remote sensing of burn severity. U. S. Geological Survey Wildland Fire Workshop, October 31–November 3, 2000. USGS Open File Report 02-11. USGS, Los Alamos, New Mexico.

Kirkman, T. P. 1847. On a problem in combinatorics. Cambridge and Dublin Mathematical Journal 2:191–204.

Kunze, L. M. 1994. Preliminary classification of native, low elevation, freshwater wetland vegetation in western Washington. Washington Natural Heritage Program, Department of Natural Resources, Olympia, Washington.

Kuramoto, R. T., and L. C. Bliss. 1970. Ecology of subalpine meadows in the Olympic Mountains, Washington. Ecological Monographs 40:317–347.

Lea, C. 2011. Vegetation classification guidelines: National Park Service Vegetation Inventory, version 2.0. Natural Resource Report NPS/NRPC/NRR—2011/374. National Park Service, Fort Collins, Colorado. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2170603 (accessed 09 December 2020).

Lea, C., and A. C. Curtis. 2010. Thematic accuracy assessment procedures: National Park Service Vegetation Inventory, version 2.0. Natural Resource Report NPS/2010/NRR—2010/204. National Park Service, Fort Collins, Colorado. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2124829.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News 2002:18–22. Available at: https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf. Code available at: https://cran.r-project.org/web/packages/randomForest/index.html (accessed 01 July 2020).

McCune, B. 2007. Improved estimates of incident radiation and heat load using non-parametric regression against topographic variables. Journal of Vegetation Science 18:751–754.

McCune, B., and D. Keon. 2002. Equations for potential annual direct incident radiation and heat load. Journal of Vegetation Science 13:603–606.

Menenti, M., and J. C. Ritchie. 1994. Estimation of effective aerodynamic roughness of Walnut Gulch watershed with laser altimeter measurements. Water Resources Research 30:1329–1337.

Moore, I. D. 1991. Digital terrain modeling: a review of hydrological, geomorphological and biological applications. Hydrological Processes 5:3–30.

Murie, A. 1935. Special report of Senior Naturalist Technician Adolph Murie, on wildlife of the Olympics. Olympic National Park, Port Angeles, Washington.

National Park Service (NPS). 2018. Vegetation mapping inventory. Available at: https://www.nps.gov/im/vegetation-inventory.htm (accessed 01 July 2020).

National Park Service (NPS). 2019. The people of the Olympic peninsula: ancient peoples and area tribes. Available at: https://www.nps.gov/olym/learn/the-people-of-the-olympic-peninsula.htm (accessed 30 October 2020).

NatureServe. 2012. NCCN alliance descriptions: Forested and a subset of non-forested alliances from Mount Rainier, North Cascades & Olympic National Parks. Interim report. NatureServe, Arlington, Virginia.

Nielsen, E. M., and R. L. Brunner. 2021. Vegetation associations for mapping Pacific Northwest national parks. Institute for Natural Resources, Portland State University, Portland, Oregon. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2283945 (available June 2021).

Nielsen, E. M., C. Copass, R. L. Brunner, and L. K. Wise. 2021a. Mount Rainier National Park vegetation classification and mapping project report. Natural Resource Report NPS/NCCN/NRR—2021/2253. National Park Service, Fort Collins, Colorado. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2286419.

Nielsen, E. M., C. Copass, R. L. Brunner, and L. K. Wise. 2021b. North Cascades National Park Complex vegetation classification and mapping project report. Natural Resource Report NPS/NCCN/NRR—2021/2254. National Park Service, Fort Collins, Colorado. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2286418.

Nielsen, E. M., R. L. Brunner, C. Copass, and L. K. Wise. 2021c. Olympic National Park map class descriptions. National Park Service, Fort Collins, Colorado. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Nielsen, E. M., C. Copass, R. L. Brunner, and K. Braun. 2021d. Olympic National Park vegetation map. National Park Service, Fort Collins, Colorado. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2284945.

Norton, H. H. 1979. The association between anthropogenic prairies and important food plants in western Washington. Northwest Anthropological Research Notes 18:175–200.

Oksanen, J., F. Guillaume Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2019. R package 'vegan': community ecology, version 2.5-6. Available at: https://cran.r-project.org/web/packages/vegan/index.html (accessed 01 July 2020).

Osborn, G., B. Menounos, C. Ryane, J. Riedel, J. J. Clague, J. Koch, D. Clark, K. Scott, and P. T. Davis. 2012. Latest Pleistocene and Holocene glacier fluctuations on Mount Baker, Washington. Quaternary Science Reviews. 49:33–51.

Pacific Meridian Resources (PMR). 1997. Vegetation and landform database development. Report to the National Park Service. Pacific Meridian Resources, Portland, Oregon.

Painter, T. H., D. F. Berisford, J. W. Boardman, K. J. Bormann, J. S. Deems, F. Gehrke, A. Hedrick, M. Joyce, R. Laidlaw, D. Marks, C. Mattmann, B. McGurk, P. Ramirez, M. Richardson, S. M. Skiles, F. C. Seidel, and A. Winstral. 2016. The Airborne Snow Observatory: fusion of scanning lidar, imaging spectrometer and physically-based modeling for mapping snow water equivalent and snow albedo. Remote Sensing of Environment 184:139–152.

Parker, G. G., and M. E. Russ. 2004. The canopy surface and stand development: assessing forest canopy structure and complexity with near-surface altimetry. Forest Ecology and Management 189:307–315.

Pelto, M. 2017. North Cascade Range, Washington USA. Pages 101–129 *in* M. Pelto. Recent climate change impacts on mountain glaciers. John Wiley & Sons Ltd., West Sussex, UK.

Pike, R. J., and S. E. Wilson. 1971. Elevation-relief ratio, hypsometric integral and geomorphic area-altitude analysis. Geological Society of America Bulletin 82:1079–1084.

Piper, C. V. 1906. Flora of the state of Washington. Contributions from the United States National Herbarium, v.11. Government Printing Office, Washington, District of Columbia.

Pojar, J., A. Mackinnon, and P. B. Alaback. 2004. Plants of the Pacific Northwest coast: Washington, Oregon, British Columbia and Alaska, second edition. Lone Pine Publishing, Vancouver, Canada.

PRISM Climate Group. 2019. PRISM climate data. Available at: https://prism.oregonstate.edu/ (accessed 01 July 2020).

R Development Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.r-project.org/ (accessed 01 July 2020).

Ramm-Granberg, T., F. J. Rocchio, C. Copass, R. Brunner and E. Nielsen. 2021. Revised vegetation classification for Mount Rainier, North Cascades, and Olympic National Parks: Descriptions and identification keys for plant associations and wetland alliances. North Coast and Cascades Network. National Park Service, Port Angeles, Washington. Available at: https://irma.nps.gov/DataStore/Reference/Profile/2279820.

Reagan, A. B. 1909. Some notes on the Olympic Peninsula, Washington. Transactions of the Kansas Academy of Science 22:131–238.

Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS. Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, volume 1A:309–317.

Smith, B. G., and J. A. Henderson. 1986. Baseline vegetation survey of the Hoh and Dosewallips drainages, Olympic National Park, Washington. Contract No. CX-9000-7-0063. Oregon State University, Corvallis, Oregon.

Therneau, T., and B. Atkinson. 2018. R package 'rpart': recursive partitioning and regression trees, version 4.1–13. Available at: https://cran.r-project.org/web/packages/rpart/index.html (accessed 01 July 2020).

Tucker, C. J., and P. J. Sellers. 1986. Satellite remote sensing of primary production. International Journal of Remote Sensing 7:1395–1416.

Twele, A. 2006. The effect of stratified topographic correction on land cover classification in tropical mountainous regions. ISPRS Commission VII Mid-term Symposium: remote sensing from pixels to processes. Enschede, The Netherlands.

U. S. Senate. 1938. H. R. 10024: an act to establish the Olympic National Park, in the State of Washington, and for other purposes. Government Printing Office, Washington, District of Columbia.

United Nations Educational Scientific, and Cultural Organization (UNESCO). 1973. International classification and mapping of vegetation. Series 6: Ecology and Conservation. UNESCO, Paris.

United States Geological Survey (USGS). 2012. National Hydrography Dataset. Available at: https://www.usgs.gov/core-science-systems/ngp/national-hydrography (accessed 01 July 2020).

United States Geological Survey (USGS). 2019a. GLOVIS: global visualization viewer. Available at: https://glovis.usgs.gov/ (accessed 01 July 2020).

United States Geological Survey (USGS). 2019b. 3D elevation program. Available at: https://www.usgs.gov/core-science-systems/ngp/3dep (accessed 01 July 2020).

United States National Vegetation Classification (USNVC). 2019. United States National Vegetation Classification Database, V2.03. Federal Geographic Data Committee, Vegetation Subcommittee, Washington District of Columbia. Available at: http://usnvc.org/ (accessed 01 July 2020).

Van Pelt, R., T. C. O'Keefe, J. J. Latterell, and R. J. Naiman. 2006. Riparian forest stand development along the Queets River in Olympic National Park, Washington. Ecological Monographs 76:277–298.

Wang, L., and H. Liu. 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. International Journal of Geographical Information Science 20:193–213.

Washington Natural Heritage Program (WNHP). 2011. Ecosystems of Washington. Available at: https://www.dnr.wa.gov/NHPecosystems (accessed 01 July 2020).

Washington State Department of Natural Resources (DNR). 2018. Pesticides and forestry fact sheet. January 2018. Olympia, Washington. Available at: https://www.dnr.wa.gov/publications/fp_pest_forestry_factsheet.pdf (accessed 01 July 2020).

Wendel, R., and D. Zabowski. 2010. Fire history within the lower Elwha River watershed, Olympic National Park, Washington. Northwest Science 84:88–97.

White, B., J. Ogilvie, D. M. H. Campbell, D. Hiltz, B. Gauthier, H. K. Chisholm, H. K. Wen, P. N. C. Murphy, and P. A. Arp. 2012. Using the cartographic depth-to-water index to locate small streams and associated wet areas across landscapes. Canadian Water Resources Journal 37:333–347.

Williams, H. 2002. The restless Northwest: a geological story. Washington State University Press, Pullman, Washington.

Wilson, E. H., and S. A. Sader. 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery. Remote Sensing of Environment 80:385–396.

Yang, L., S. Jin, P. Danielson, C. Homer, L. Gass, S. M. Bender, A. Case, C. Costello, J. Dewitz, J. Fry, M. Funk, B. Granneman, G. C. Liknes, M. Rigge, and G. Xian. 2018. A new generation of the United States National Land Cover Database: requirements, research priorities, design, and implementation strategies. ISPRS Journal of Photogrammetry and Remote Sensing 146:108–123. Available at: https://www.mrlc.gov/data/nlcd-2016-land-cover-conus (accessed 01 July 2020).

The Department of the Interior protects and manages the nation's natural resources and cultural heritage; provides scientific and other information about those resources; and honors its special responsibilities to American Indians, Alaska Natives, and affiliated Island Communities.

NPS 149/175302, May 2021

Natural Resource Stewardship and Science
1201 Oakridge Drive, Suite 150
Fort Collins, CO 80525