# Guide to Understanding and Using the Potential Natural Vegetation (PNV) Map Layer

(June 2022)

The purpose of this guide is to introduce the user to the concept of Potential Natural Vegetation (PNV), the new PNV maps available at vegetation zone and subzone scales, and applications of the maps to address resource assessment and planning needs. This guide will accompany the PNV layer. A general technical report is currently underway to delve deeper into the topics covered within this guide.

## Background PNV concepts

PNV is a classification system that uses indicator plant species to designate a biophysical environment (ecological potential) and infer a potential suite of ecosystem structures and compositions, along with functions, including disturbance, through time. It integrates the climate, geology, geomorphology, soils, and vegetation in an area. PNV is a biological indicator of land capability for supplying the array of ecosystem services the landscape provides.

The Forest Service has defined PNV (from Winthers et al. 2005) as
.. *the vegetation that would become established if all successional sequences were completed without human interference under present climatic and edaphic conditions*
[adapted from Tüxen (1956) as translated by Mueller-Dombois and Ellenberg (1974)].

We define PNV with a more contemporary focus on describing the biophysical environment, incorporating the importance of ecological characteristics and processes (notably disturbance), across landscapes. In this view, potential vegetation can be used to infer productivity, frame disturbance regimes and expected response to disturbance, describe wildlife habitat potential, and at broader scales, provide a coarse filter approach to assessing biodiversity and landscape pattern. PNV informs what management is both possible and recommended for desired conditions, the consequences of management actions, and what ecological goods and services can be generated, apart from the current existing vegetation.

PNV names are useful for describing site environmental conditions that are relatively stable through time and relating them to vegetation that can be observed on the ground. These names can be expressed at various scales. Warm, dry environments labeled by the dominant indicator tree species, such as “ponderosa pine”, implies a climate, productivity, fire regime, and wildlife habitat suite much different than that of moist, warm to cool “western hemlock”. Presence and cover of other indicator tree, shrub, herb, and graminoid species can help determine biophysical environments at a finer scale, describing the variability that exists within a broader type. Sites in the western hemlock type, where indicator species such as salmonberry common ladyfern, deer fern, devilsclub, and false lily of the valley occur, identify conditions where water collects, often adjacent to riparian areas and/or poorly drained sites. When indicator species such as dwarf Oregongrape, common prince’s-pine, oceanspray, starflower, vanillaleaf, western rattlesnake plantain, western larch, and common whipplevine occur, one can infer mesic site conditions and the driest environments in the western hemlock type.

PNV is not interchangeable with “desired” or “sustainable” conditions. With a changing climate, the potential vegetation of the past may no longer be attainable. With a greater appreciation of the role of disturbance (notably fire but also insect outbreaks and other disturbances) in recent years, ecologists question the very idea of plant community stability. In response to shifting climate and disturbance regimes, previously associated species may migrate at different rates, forming new communities that do not fit into current PNV categories. Even with these constraints, PNV reflects a substantial body of knowledge about the relationships between vegetation, land, and ecosystem processes, and provides a baseline from which to consider the effects of global ecological change and local management (Somodi et al. 2012).

## Previous R6 PNV classification and mapping efforts

Ecologists have classified PNV at various scales to meet management needs on federal lands in Washington and Oregon. Data on plant species abundance in late seral areas, along with site data (slope, elevation, aspect, soil, etc.) was collected by ecologists to develop the classification. Environmental indicator tree, shrub, herb, and graminoid plant species were identified and used to define **plant associations**, named by climax plant communities that exhibited repeating patterns of vegetation composition, associated with repeating patterns of climate, soil, and topography (McCune and Grace 2002). In forest and shrubland systems, overstory indicator species have been used to separate broad types of vegetation; these groupings are called **series** (e.g., Western Hemlock, Ponderosa Pine, Mountain Big Sagebrush). Understory indicator species (e.g., dwarf Oregongrape, baldhip rose, common prince’s-pine, and California blackberry for the Western Hemlock Series; antelope bitterbrush, squirreltail, Idaho fescue, and western juniper for the Ponderosa Pine Series), provide information on local site conditions to further refine biophysical environments into plant associations/habitat types. The Region 6 plant association guides are found on the website ECOSHARE (https://ecoshare.info/).

The PNV classification has proved useful in a variety of management arenas. Once the concepts began to be implemented there was a strong desire to be able to capture biophysical environment distribution across the landscape. Shifts in environmental factors causing changes in plant associations can occur at a very fine scale, and environmental distribution data at the micro-scale is currently not available for mapping. As a result, plant associations/habitat types were aggregated into plant association/habitat type groups (PAGs) that shared similar dominant vegetation, environmental setting factors, and/or temperature and moisture gradients that could be mapped across a landscape.

PAG mapping began in the 1990s-2000s. Although many methods were initiated (e.g., on the ground mapping, aerial photo interpretation), they were abandoned due to intensive labor requirements and varying results. Former Area Ecologist for western Washington, Jan Henderson, along with Ecologists Robin Lesher, David Peter, and Chris Ringo, developed a potential vegetation mapping process using polynomial equations tying environmental gradients to ground data to generate a potential vegetation map (Henderson et al. 2011). This method was employed across Washington state, and in some areas in Oregon, but was never fully implemented across the Region. Mapping efforts also struggled with edge matching issues, making the utility limited for more regional needs.

Another mapping effort, the Integrated Landscape Assessment Project (ILAP), produced consistent, landscape-wide vegetation mapping across Oregon and Washington to assess current conditions, expected trends, and possible implications of alternative policies and management actions. The potential vegetation component used to stratify the existing vegetation data for departure assessment used a combination of each of the Area's potential vegetation mapping products. These products were not all created using similar methods. In addition, the differences between each area potential vegetation classification created some problems with departure interpretations across area boundaries and may have limited utility for rollup to regional scale departure assessments. This mapping effort was used extensively within the DecAID program (Decayed Wood Advisor, https://apps.fs.usda.gov/r6_decaid/).

There have been other national attempts to describe and map PNV. The National Landfire Project (https://landfire.cr.usgs.gov/) defines PNV using the “disturbance-constrained” approach. PNV types are named after the vegetation community that would be expected on a site under its characteristic disturbance regime. For example, frequent fire may have historically maintained open ponderosa pine forest on a site, but Douglas-fir would be the climax tree component under the site’s climate, topography, and soil conditions. The climate-constrained PNV convention used in R6 might name these landscapes as “dry Douglas-fir”, whereas the disturbance-constrained convention would name them “ponderosa pine”. Both names are shorthand for a range of possible vegetation states on that site. Douglas-fir may currently dominate this environment as the result of a century of extractive land use and fire exclusion, and the loss of ponderosa pine in these forests may have negative ramifications for future forest function, wildlife habitat considerations, and fire risk. In this case, knowledge of both climate- and disturbance-constrained PNV in a given area is critical for assessing ecosystem integrity, understanding processes that favor one vegetation state over another, and then for identifying possible management alternatives. Crosswalks between the two approaches are available, however, the crosswalks are not truly a one-to-one relationship. In fact, in many cases they are many to many relationships which creates many interpretation challenges.

## Need and uses for consistent classification and mapping

PNV work in the Forest Service in the 1970s and 1980s originally developed around the need to match tree seedlings to site, and quickly grew to include forest productivity, wildlife habitat, disturbance regime, and management response issues. Existing vegetation seral stages can be associated with each of the landscape PNV types, to facilitate classification and mapping on topics such as late seral forest, old growth, and the need for quality early seral forest. Since many of these issues are regional in scale, the need for a consistent approach in both classification and mapping of PNV is essential.

One example of using PNV involves organizing PNV into wildlife habitat potential through space and time. Snags and down wood are important components of ecosystems and a major wildlife concern. Forest types in DecAID (a decaying wood advisory model for Oregon and Washington), a regional product, are organized by wildlife habitat types. Information in DecAID is organized by wildlife habitat types, which are intended to capture broad classes of current vegetation with similar patterns of wildlife use (O'Neil, T.A., Johnson, D.H., 2001). As a practical matter, wildlife habitat types are mapped based on a crosswalk to potential vegetation, because PNV provides a good metric on biomass productivity of ecosystems. Better use of landscape scale consistent PNV maps, coupled with information on existing vegetation structure and successional pathways, make for a powerful wildlife habitat planning tool, and facilitates a more complete understanding of habitat.

Potential vegetation types can also be associated with fire regimes and fire effects because each PNV type has a characteristic topography, climate, and productivity, along with a list of common species found on sites through time. Information can be gleaned on potential response to fire, with coupling of PNV and current condition. PNV Groups may also be used to further refine fire regimes. If a particular, smaller in area, PNV Group is surrounded by a large patch of a different PNV Group with a different fire regime, it may behave during a disturbance/fire more like the dominant type that it is surrounded by, or adjacent to, than its own fire regime.

PNV can also provide a critical framework for organizing the landscape for ecological departure assessment, an important measure of landscape resilience and resistance. To our knowledge, all papers on departure from a natural range of variation (NRV) landscape assessments use some form of potential vegetation to organize landscapes (Columbia Basin work, Barrett et al. 2010, Haugo et al. 2015, DeMeo et al. 2018, Washington DNR work, Olszewski et al. in press, Hagmann et al. in press, cite others). Our work on this PNV mapping effort has been motivated in part to provide refined map units to support NRV analysis given the need for consistent, broad scale information.

PNV can also provide critical information on targeting response to treatments, both for commodity production and restoration. Consistent PNV landscape scale classification and maps can help address questions such as:

- Where are the most productive forests for timber production, e.g., biomass, height, and timber volume?
- What species may be present, or most adapted to the landscape through time, given a changing climate?
- What is the land suitability for planting trees, by species, and the likelihood of success?
- What will various thinning, and/or prescribed burning treatment responses be (understory and overstory)?

- Where is thinning and prescribed burning most needed for restoration purposes?
- Where are invasive species most likely to be a problem?
- What is the most sustainable mix of seral stages, and how much of each is needed?
- What are treatment responses to fuels reduction, and where may favorable effects be most successful?

Climate change, and its effects across the landscape, is another aspect where PNV can be useful. Climate change is affecting landscapes across the Pacific Northwest, trending towards warmer temperatures and more irregular patterns in precipitation. PNV mapping provides a useful tool to show how climate change effects vary with location and vegetation zone. Some issues relevant to this are higher elevation snowpack, drought patterns, and effects on the size and severity of wildfires. PNV maps can be used to identify climate sensitive forests, organize the magnitude of the effects of climate change, and imply what can be done to address them.

All these aspects of the utility of consistent PNV classification mapping at broad scales, such as framing disturbance regimes, organizing wildlife habitat, understanding productivity, planning for restoration, and understanding climate change effects, make this foundational suite of products an essential component of the forest planning process. For example, potential vegetation maps were used in the Bioregional Assessment conducted for the Northwest Plan area (https://www.fs.usda.gov/detail/r6/landmanagement/planning/?cid=fseprd677501), the revision of Eastside screens direction (https://www.fs.usda.gov/detail/r6/landmanagement/planning/?cid=FSEPRD731318), Northwest forest plan northern spotted owl monitoring reports (Davis et al 2016; Davis et al 2021), and for the current Forest Plan revision process for the five National Forests in the Klamath Province of Northern California and Southwest Oregon.

## Methods

The two-tiered approach we used to classify, and map PNV, was similar to previous classification efforts. The **vegzone** is synonymous with the **series** and uses the indicator overstory (mainly tree) species. The **subzone** is most like, although generally broader, the **plant association group,** and uses key indicator tree, shrub, herb, and graminoid species to identify the range of finer scale biophysical environments that fall within each vegzone.

We built the new PNV map on the foundation of the wall-to-wall, 30-meter maps of current forest attributes developed using the gradient nearest neighbor (GNN) imputation mapping approach (Bell et al. 2021). GNN integrates Landsat multispectral remote sensing data, forest inventory plot data, and environmental data to assign current compositional and structural attributes from field observations to all land area deemed to have the potential to support forest (Ohmann and Gregory 2002, Bell et al. 2021). GNN provides some desirable qualities for mapping potential natural vegetation due to its spatially comprehensive nature, its temporal depth (data available from 1986-2017), and its inherent incorporation of environmental drivers of spatial patterns of primary productivity and species composition.

## Classification rules and ranking for both vegzone and subzone

We exploited species composition data from the inventory plots used for GNN to assign each plot to a potential vegzone and subzone. For **vegzone** assignments, we established a hierarchical ranking of those tree species used as indicator species in plant association guides for National Forests across the study area, taking into account shade tolerance and longevity of species, and adhering as much as possible to the logic in the various plant association guides. Shade-tolerant species with the narrowest environmental distribution (e.g., Sitka spruce, redwood, and mountain hemlock) are highest in the hierarchy, followed by shade-tolerant species of broader distribution (e.g., Pacific silver fir, western hemlock), followed by less shade-tolerant species that represent earlier-seral conditions in many environments (e.g., Douglas-fir, lodgepole pine). We used a cover threshold of 5% for roughly half of the indicator tree species, with most others having thresholds of 10% or 20%; for about one-third of indicator species, cover thresholds were relaxed for those inventory plots judged to have been recently disturbed (i.e., stand age of less than 80 years). In the case of Sitka spruce, the cover threshold was 10% for undisturbed plots while the threshold for disturbed plots was presence of the species. The complete rule set included many additional contingencies based on the abundance of co-occurring species, other than the named indicator species, and will be included in the general technical report. It is also available upon request.

To assign **subzones** to inventory plots, we considered indicator graminoid, herb, shrub, and tree species not used in vegzone assignments, primarily those that were used in the plant association guides. Using inventory plots and the ecology plots (which were used in developing the plant association guides), we examined species distributions with respect to climatic variables and their affinities to one another in species-composition space. We established nine species groups indicating different combinations of temperature and moisture regimes. These were ordered hierarchically from most productive (e.g., warm and wet) to least productive (xeric). In most cases, we used a threshold of one species and 25% cover or 10% combined cover and more than one species in the environmental group, to assign plots to subzones. As for vegzones, a plot location assigned different subzones over time was labeled for the subzone highest in the hierarchy. We also established a separate list of species that indicated serpentine substrates.

To develop consistent terminology for subzone names, we used distributed environmental data in conjunction with vegetation data. We summarized environmental data by vegzone and subzone for the ecology plots as a basis for defining subzones. For most vegzone/subzone combinations, we were able to assign labels using values of environmental variables associated with ecology plots (R6 Ecology program). For a small number of combinations (mostly those found in California and not in Oregon or Washington), we used environmental data associated with pixels in the new map representing the relevant zone/subzone combinations.

For both types of vegetation data, we applied a consistent set of descriptors linked to median values of mean annual precipitation and mean annual temperature (Tables 1 and 2, respectively).

Table 1. Descriptor terms based on the median of mean annual precipitation (inches) for vegetation zone/subzone combinations in new map of potential natural vegetation.

| Lower limit (≥) | Upper limit (<) | Descriptor term |
|---|---|---|
| 80 | n/a | VeryWet |
| 70 | 80 | Wet |
| 50 | 70 | VeryMoist |
| 30 | 50 | Moist |
| 20 | 30 | Dry |
| n/a | 20 | Xeric |

Table 2. Descriptor terms based on the median of mean annual temperature (°F) for vegetation zone/subzone combinations in new map of potential natural vegetation.

| Lower limit (≥) | Upper limit (<) | Descriptor term |
|---|---|---|
| 47.5 | n/a | VeryWarm |
| 45 | 47.5 | Warm |
| 43 | 45 | Mild |
| 41 | 43 | Cool |
| n/a | 41 | Cold |

In cases where more than one subset of a vegzone was assigned the same combination of annual precipitation and annual temperature descriptors, we exploited other environmental variables (e.g., seasonal temperatures, slope position) to identify descriptors to distinguish between the subzones. For vegetation zone/subzone combinations represented in the Ecology plots, we compared inter-quartile ranges for precipitation and temperature. For vegetation zone/subzone combinations not in the Ecology plots but evaluated using locations in the potential vegetation map, we compared mean values plus and minus 1.2 times the standard deviation (also representing 75% of observations for normally-distributed data; see https://en.wikipedia.org/wiki/Standard_deviation). In the case of Pinyon-Juniper-Cypress Xeric Warm, no available environmental variables were useful in distinguishing the ambiguous groups. In these cases, we used labels derived from physiognomy, composition, or habitat.

## Temporal stack

We used repeated field observations corresponding to the entire temporal span of GNN (1986-2017) to maximize the chance of capturing later-seral species for each plot location. These repeated measurements allowed us to understand and accommodate successional stages. We evaluated the assigned **vegzone** for each year of plot measurements at a particular location, selecting the vegzone that represented the highest level in the hierarchy (for example, a location with Western Hemlock and Douglas-fir Vegzone assignments at different years would be labeled as the Western Hemlock Vegzone). This process relied only on the inventory data and was completely independent of mapped GNN predictions.

With all inventory plot locations for GNN labeled with a single vegzone and subzone, it was possible to generate provisional maps of PNV for each year in the entire temporal span of GNN. In a GIS exercise, we then determined the majority, variety, and highest hierarchical rank for each pixel in the study area. Where fewer than or equal to four subzone types were assigned to a pixel over time, the final label was the majority subzone. Where greater than four subzone types were assigned to a pixel over time (suggesting the location had experienced disturbance), or there was no majority subzone (i.e., ties), the final label was the highest hierarchical rank.

## Non-Forest Mask Development

PNV types other than forest were added to the vegzone and subzone maps based on existing non-forest mapping methods, starting with the 2012 GNN non-forest mask (see https://lemma.forestry.oregonstate.edu/). We modified this mask to reduce embedded disturbance artifacts (mostly old fire scars) and to improve the wetlands delineations within the mask. We used a combination of USGS National Land Cover database, 1992 and 2011 USGS continuous forest cover estimates, and GNN maximum forest cover predictions from 1984-2017 as the basis for this analysis with the assumption that pixels with >= 20% tree cover at any point in time in any data source are forested or woodland types. Where LiDAR was available, we also used LiDAR data and assumed locations (30-meter pixels) with >=10% LiDAR first returns above 3.5 meters have previously supported forest and therefore have a forested potential vegetation type. The USGS wetlands inventory (NWI) layer was used to identify and categorize emergent and palustrine wetlands.

The USGS-Landfire ESP (Ecological Systems) layer was used to populate the vegetation zones and sub-zones within the non-forest mask. The Ecological Systems were grouped into functional groups for the vegzone layer and include Parklands, Shrublands, Grasslands-Meadows, Ice and Snowfields, Rock, Water, and Developed. Subzones were created for the non-forest portions of Parklands, as well as for the Shrublands, and Grasslands-Meadows using groups of combinations from both the USGS-Landfire ESP and USGS NWI wetland classifications. The final map was filtered twice (majority call based on a 3x3 grid to smooth the map) and spot checked using historical photomosaic data.

## Accuracy assessment

We took a three-pronged approach to assessing accuracy of the PNV map using the “nearest neighbor” concepts from GNN:

- **Model stability/uncertainty** – check the similarity among vegzone/subzone rasters using the second, third, etc. neighbors rather than the first neighbor.
- **Local scale accuracy** - Observed vs. predicted vegzone/subzone by location
- **Area based regional accuracy** - Design-based (plot) vs. model-based (GNN) area estimates

### Model stability/uncertainty

We compared vegzone/subzone-only realizations using the second, third, fourth and fifth nearest neighbors (k=2, 3, 4, 5) to the observed subzone (first nearest neighbor (k=1)) to assess how stable the model was. We created maps using the other neighbors for each year in the 1986-2017 temporal stack and followed the same rules in our methods section to create a single subzone raster for each of the neighbors. In general, exact matches to our PNV map (k=1) decline when using the other nearest neighbors. The comparisons vary by subzone, but typical declines in accuracy from k1k2 to k1k5 are 10-30% with an average decline in overall accuracy of 19%.

We also used the cell statistics tool in ArcGIS 10.5 to evaluate map stability by looking at the cell majority (the value most common when looking across all the temporal maps), cell variety (how many different subzone calls were made for a pixel over the temporal maps), and min-rank (the vegzone call that was the highest in the rule hierarchy) for each 30-meter cell across all the five nearest neighbor subzone maps. We then combined the k1, cell majority, cell variety, and min-rank layers into a single multiband raster (f_k1_k1k5_comb). We used the cell variety raster to look at agreement across all k1-k5 realizations. Variety values 1-2 (i.e., 1 – 2 distinct subzones predicted across five realizations) (78%) have good agreement across all k1-k5 rasters while values 3-4 (22%) demonstrate areas in the map where the modeled subzone is more uncertain. The uncertain areas appear to be a combination of disturbed areas and transitional areas from one vegzone to another.

Two additional rasters were built from the f_k1_k1k5_comb raster and each was evaluated for exact and fuzzy match. The first raster compared k1 subzone to k1k5 subzone majority, the second raster compared k1 to k1k5 subzone majority for k1k5 variety 1-2 and k1k5 subzone minimum rank for k1k5 variety 3-4. The second k1k5 comparison was built to emulate the subzone aggregation logic which utilizes the minimum ranked subzone in areas where there is higher uncertainty. The k1 to k1k5 majority raster and the k1 to k1k5 majority/minimum rank raster have similar exact match and fuzzy match agreements, however the k1 to k1k5 majority raster has 4-5% higher agreement (Figure 1). The highest exact match that we could expect across the whole map area is around 68% (fuzzy is 90%), partially due to small sample sizes associated with some subzones.

Figure 1 – Comparison of exact and fuzzy matches between subzone (k1) to k1_k1k5_maj and k1_k1k5_majmin rasters.

| | k1_k1k5_maj | k1_k1k5_majmin |
|---|---|---|
| Exact_Match | 68% | 63% |
| Fuzzy_Match | 90% | 86% |

**Local scale accuracy**

We used a "nearest-independent neighbor" approach to assess accuracy of the new map for both vegzone and subzone (Ohmann and Gregory, 2002.). The nearest-independent neighbor approach allows us to generate a model validation dataset that behaves similarly to an independent validation dataset from the same forest inventory data used to develop GNN maps, thus providing a more reliable assessment of PNV map accuracy. For each inventory plot location used in producing the GNN dataset, the nearest neighbor plots in gradient space that are not measured at the same physical location were selected for the nine 30 X 30-meter pixels that comprise the areal extent of the plot's sample area. In this way we produced a time-series of 32 (1986-2017) nearest-neighbor inventory plots for each pixel in a target-plot footprint. These plots were all assigned a subzone using the same logic as for the new PNV map. To get to a single subzone assignment for each pixel, we evaluated the majority (mode) subzone across all 32 observations and the number of distinct subzone assignments across all 32 observations. When the number of distinct subzone assignments was <=4, we assigned the majority subclass. Otherwise, we assigned the subzone the highest hierarchical rank, or minimum rank. Finally, we took the majority (mode) across all nine pixels in the location footprint to represent the predicted subzone. To assign predicted vegetation zones, we applied the vegetation zone associated with each predicted subzone.

For both vegetation zones and subzones, we performed both a "strict" and "fuzzy" version of the accuracy assessment. In the strict version, an assignment was recorded as correct if there was an exact match between the original assignment and the predicted assignment using the nearest-independent neighbor method. To carry out the fuzzy version, we constructed lists for each subzone of other subzones sufficiently similar in their environmental relationships and floristic composition that interpretation of the map would not significantly suffer if the assignments were swapped. For example, for the vegetation zone represented by the largest number of unique inventory plot locations (White Fir-Grand Fir), there were 11 vegetation zones that we judged to constitute a fuzzy match: Lodgepole Pine, Ponderosa Pine, Jeffrey Pine-Knobcone Pine, Douglas-Fir, Giant Sequoia, Western Hemlock, Western Red Cedar, California Red Fir-Shasta Red Fir, Pacific Silver Fir, Mountain Hemlock, and Subalpine Fir-Engelmann Spruce. The strict producer's accuracy (i.e., percentage of plots in the zone that were mapped correctly) for the White Fir-Grand Fir Vegzone was 71%, whereas the fuzzy producer's accuracy was 99%. The strict user's accuracy (i.e., percentage of the plots mapped as the zone that were observed in that zone) for the White Fir-Grand Fir zone was 71%, whereas the fuzzy user's accuracy was 98%.

For vegetation zones, the overall strict accuracy was 65% and the overall fuzzy accuracy was 97%. For individual zones, strict producer's accuracy ranged from 0% to 89% and fuzzy producer's accuracy ranged from 0% to 100%, while strict user's accuracy ranged from 0% to 78% and fuzzy user's accuracy ranged from 0% to 100%. The lower values are mostly associated with very minimally represented zones (i.e., Port Orford Cedar, Giant Sequoia, Grasslands-Meadows, and Shrublands, each of which accounted for at most 0.1% of the map). For the 10 most common vegetation zones, which together accounted for 87% of the map (White Fir - Grand Fir, Western Hemlock, Douglas-Fir, Ponderosa Pine, Pacific Silver Fir, Pinyon-Juniper-Cypress, Mountain Hemlock, Hardwoods, Subalpine Fir - Engelmann Spruce, and Tanoak), strict producer's accuracy ranged from 56% to 82% and fuzzy producer's accuracy ranged from 90% to 99%, while strict user's accuracy ranged from 51% to 78% and fuzzy user's accuracy ranged from 92% to 100%.

For subzones, the overall strict accuracy was 42% and the overall fuzzy accuracy was 85%. For individual subzones, strict producer's accuracy ranged from 0% to 88% and fuzzy producer's accuracy ranged from 0% to 100%, while both strict and fuzzy user's accuracy ranged from 0% to 100%. The lower values are associated with very minimally represented subzones (i.e., Moist Western Hemlock, Limber Pine Parklands, Cool Western Red Cedar, White Bark Pine Parklands, and the one Giant Sequoia subzone, each of which accounted for at most 0.2% of the map). For the Western Hemlock VeryMoist Subzone, the most common subzone, strict producer's accuracy was 71%, fuzzy producer's accuracy was 94%, strict user's accuracy 63%, and fuzzy user's accuracy was 91%.

### Area based regional accuracy

We compared designed-based (FIA inventory plots) and model-based (using GNN) area estimates for each vegzone and subzone. We used the error matrix from the local scale accuracy test to adjust model-based (GNN) area estimates and generate adjusted area estimates plus 95% confidence intervals based on methods described by Olofsson et al., 2013.

To determine agreement between the two area estimates, we employed the following methods. For design-based estimates we used species composition from the 2016 FIA annual inventory to determine both subzone and vegzone. Design-based plots were expanded by their expansion factors. For GNN, we do this in a much simpler way than official FIA methods. We simply use the number of plots in each of our strata (Washington, Oregon, California federal lands, California non-federal lands) and divide total area by this number. This is done for forested and non-forested plots together for Oregon and Washington. Furthermore, each plot's area expansion factor is multiplied by the proportion of forest/non-forest/non-sampled on that plot such that it is divided proportionally on multi-condition-class plots. Plots were then grouped by like vegzone and subzone and area represented by the plots in each classification unit were summarized. For model-based estimates we summarized the pixels in each vegzone and subzone from our final map product. We calculated the error-adjusted area estimates by using the error matrix from the observed/predicted local scale assessment to adjust map-based areas toward plot-based proportions. For example, assume we have a large over-prediction of a certain GNN class relative to what shows up in the design-based sample. Only if the error matrix shows this large proportional over-prediction (i.e., commission) will the Olofsson area-estimate be adjusted to a lower area estimate. If errors are proportionally balanced in the error matrix, the corrected area estimate will be roughly the same as the mapped estimate as we are assuming that the errors, which we see at plots, apply everywhere.

The error correction highlights which classification units were over predicted or under predicted. Results indicated a general close match between design and model-based estimates for vegzone. 53% of the individual vegzones were within the Olofsson 95% confidence interval. Modeled area estimates averaged 96% within the Olofsson 95% confidence interval of the design-based area estimate and ranged from 86 – 100% for individual vegzones. The following vegzones model-based area were overestimated: Pinyon-Juniper-Cypress, Foothill Pine-Coulter Pine, Sitka Spruce, Redwood, Western Hemlock, and Mountain Hemlock, while the following vegzones model-based area were underestimated: Lodgepole Pine, Ponderosa Pine, Douglas-fir, and Parklands.

Subzone results were similar to the vegzone results indicating a general close match between design and model-based estimates. 71% of the individual subzones were within the Olofsson 95% confidence interval. Modeled area estimates averaged 97% within the Olofsson 95% confidence interval of the design-based area estimate and ranged from 56 – 100% for individual subzones. The following subzones model-based area were overestimated: Oak Woodlands, Juniper Woodlands, Foothill Pine-Oak, VeryWet Sitka Spruce, VeryMoist Redwood, VeryMoist Western Hemlock, CoolWet Mountain Hemlock, and Subalpine Larch Parklands, while the following subzones model-based area were underestimated: Riparian Hardwood Forest, Other Hardwoods, CoolDry Lodgepole Pine, WarmXeric Ponderosa Pine, VeryWarmDry Douglas-fir, WarmDry Douglas-fir, VeryWarmMoist Douglas-fir, CoolMoist White Fir-Grand Fir, Moist Western Hemlock, and Whitebark Pine Parklands.

## Appropriate scales

Scale is essential to understanding patterns and processes across the landscape, with the goal to apply this understanding appropriately in management. PNV can be applied at multiple scales, from local site tree regeneration to landscape disturbance processes, with various resource management implications, such as silviculture, wildlife habitat, soils, and fire ecology. PNV can also be used to describe or frame ecological structure, e.g., characteristic tree diameters, canopy layers, biomass, function (disturbance processes, nutrient regimes, and hydrologic processes), and composition (absolute and relative abundances of plant species), at multiple scales. Some questions are best answered at a particular scale. While some ecological characteristics, such as fire/disturbance regime, only emerge at broader scales, some, such as sub-watershed or individual stand treatment objectives, occur at smaller scales. This PNV classification and mapping effort was designed to be used at regional and sub-regional scales, although it has proven to have mid-scale possibilities, given the accuracy assessment results, especially in areas of high local and area agreement. In addition, Bell et al. (2018) found that GNN based above ground biomass estimates stabilized, compared to LiDAR based above ground estimates, at aggregate scales between 10-100 hectares (roughly stand scales). While

GNN and LiDAR highlighted similar patterns of high versus low biomass forests, biases between the two maps were still common. The consistent classification process employed for this effort makes this product useful across broad landscapes, however, in some areas, site specific environments may have filtered out what the plant association guides unearthed. Therefore, due to the inherent uncertainties in the PNV map, especially at finer scales, local evaluation is advised.

Table 3 outlines potential applications of the available PNV mapping efforts.

Table 3. How potential natural vegetation (PNV) varies with scale, drivers, and applications.

| **Mapping** | **Classification** | **Drivers** | **Example Applications** |
|---|---|---|---|
| Vegetation Zone | Series or groups of Series | Geology, Geomorphology, Climate, Floristics | Broad planning framework, disturbance (fire regime), wildlife habitat |
| Subzone | Subseries or groups of Subseries | Geology, Geomorphology, Climate, Floristics | Broad planning framework, disturbance (fire regime); Mid-scale framework (disturbance, treatment types and response) |
| Plant Association Group (PAG)/ Habitat Type Group (HTG) | Subseries | Climate, Soils, Floristics (species composition) | Local planning framework, silvicultural prescriptions |
| Local Plant Association | Plant associations/Habitat types | Climate, Floristics (species composition) | Local planning framework, silvicultural prescriptions |

## 5.0 References

Barrett, S., Havlina, D., Jones, J., Hann, W.J., Frame, C., Hamilton, D., Schon, K., DeMeo, T., Hutter, L., Menakis, J., 2010. Interagency Fire Regime Condition Class (FRCC) Guidebook, version 3.0. In: USDA Forest Service, US Department of the Interior, and The Nature Conservancy. <http://www.frcc.gov/>.

Bell, D. M., Acker, S. A., Gregory, M. J., Davis, R. J., & Garcia, B. A. (2021). Quantifying regional trends in large live tree and snag availability in support of forest management. *Forest Ecology and Management*, *479*, 118554.

Bell, D. M., Gregory, M. J., Kane, V., Kane, J., Kennedy, R. E., Roberts, H. M., & Yang, Z. (2018). Multiscale divergence between Landsat-and lidar-based biomass mapping is related to regional variation in canopy cover and composition. *Carbon balance and management*, *13*(1), 1-14.

DeMeo, T., Haugo, R., Ringo, C., Kertis, J., Acker, S., Simpson, M., & Stern, M. (2018). Expanding our understanding of forest structural restoration needs in the Pacific Northwest. *Northwest Science*, *92*(1), 18-35.

ECOSHARE (n.d.). Interagency Clearinghouse of Ecological Information. https://ecoshare.info/

Haugo, R., Zanger, C., DeMeo, T., Ringo, C., Shlisky, A., Blankenship, K., ... & Stern, M. (2015). A new approach to evaluate forest structure restoration needs across Oregon and Washington, USA. *Forest Ecology and Management*, *335*, 37-50.

Henderson, J. A., Lesher, R. D., Peter, D. H., & Ringo, C. D. (2011). A landscape model for predicting potential natural vegetation of the Olympic Peninsula USA using boundary equations and newly developed environmental variables. *Gen. Tech. Rep. PNW-GTR-841. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 35 p*, *841*.

Hessburg, P. F., Charnley, S., Wendel, K. L., White, E. M., Singleton, P. H., Peterson, D. W., ... & White, R. (2020). The 1994 Eastside Screens large-tree harvest limit: review of science relevant to forest planning 25 years later. *Gen. Tech. Rep. PNW-GTR-990. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 114 p.*, *990*.

Hessburg, P., Salter, R. B., Richmond, M. B., & Smith, B. G. (2000). Ecological subregions of the interior Columbia Basin, USA. *Applied Vegetation Science*, *3*(2), 163-180.

Null 1999. The Interior Columbia Basin Ecosystem Management Project: scientific assessment. CD-ROM. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

LANDFIRE (n.d.). Landscape Fire and Resource Management Planning Tools. https://landfire.cr.usgs.gov/

McCune, B., Grace, J. B., & Urban, D. L. (2002). *Analysis of ecological communities* (Vol. 28). Gleneden Beach, OR: MjM software design.

Mueller Dombois, D., & Ellenberg, H. (1974). *Aims and methods of vegetation ecology* (No. 581.5 M8).

Ohmann, J. L., & Gregory, M. J. (2002). Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research*, *32*(4), 725-741.

O'Neil, T.A., Johnson, D.H., 2001. Oregon and Washington wildlife species and their habitats. In: Johnson, D.H., O'Neil, T.A. (Eds.), Wildlife-habitat relationships in Oregon and Washington. Oregon State University Press, Corvallis, OR, USA, pp. 1–21.

Somodi, I., Molnár, Z., & Ewald, J. (2012). Towards a more transparent use of the potential natural vegetation concept–an answer to Chiarucci et al. *Journal of Vegetation Science*, *23*(3), 590-595.

USDA Forest Service, Regions 5 and 6. 2020. *Bioregional assessment of Northwest forests*. USDA Forest Service. 89 p.

Winthers, E., Fallon, D., Haglund, J., DeMeo, T., Nowacki, G., Tart, D., ... & Robbie, W. (2005). Terrestrial ecological unit inventory technical guide. *Gen. Tech. Rep. WO-GTR-68. Washington, DC: US Department of Agriculture, Forest Service, Washington Office, Ecosystem Management Coordination Staff. 245 p.*, *68*.