Comparison of land use/land cover for sites from random sampling designs (random) versus other site selection methods (non-random)

Jo Wilhelm, King County, April 4, 2012

Snohomish and King Counties dominate the site selection regardless of whether random or all sites are used (Figure 1, green points are random, red are non-random). However, if just random sites are selected, we lose quite a bit of coverage in Clallam, Pierce, Thurston, and Whatcom counties. Jefferson County has no sites with data in the PSSB; Skagit County only has 3 sites and all are random; Mason County only has 6 sites total, three of which are random.

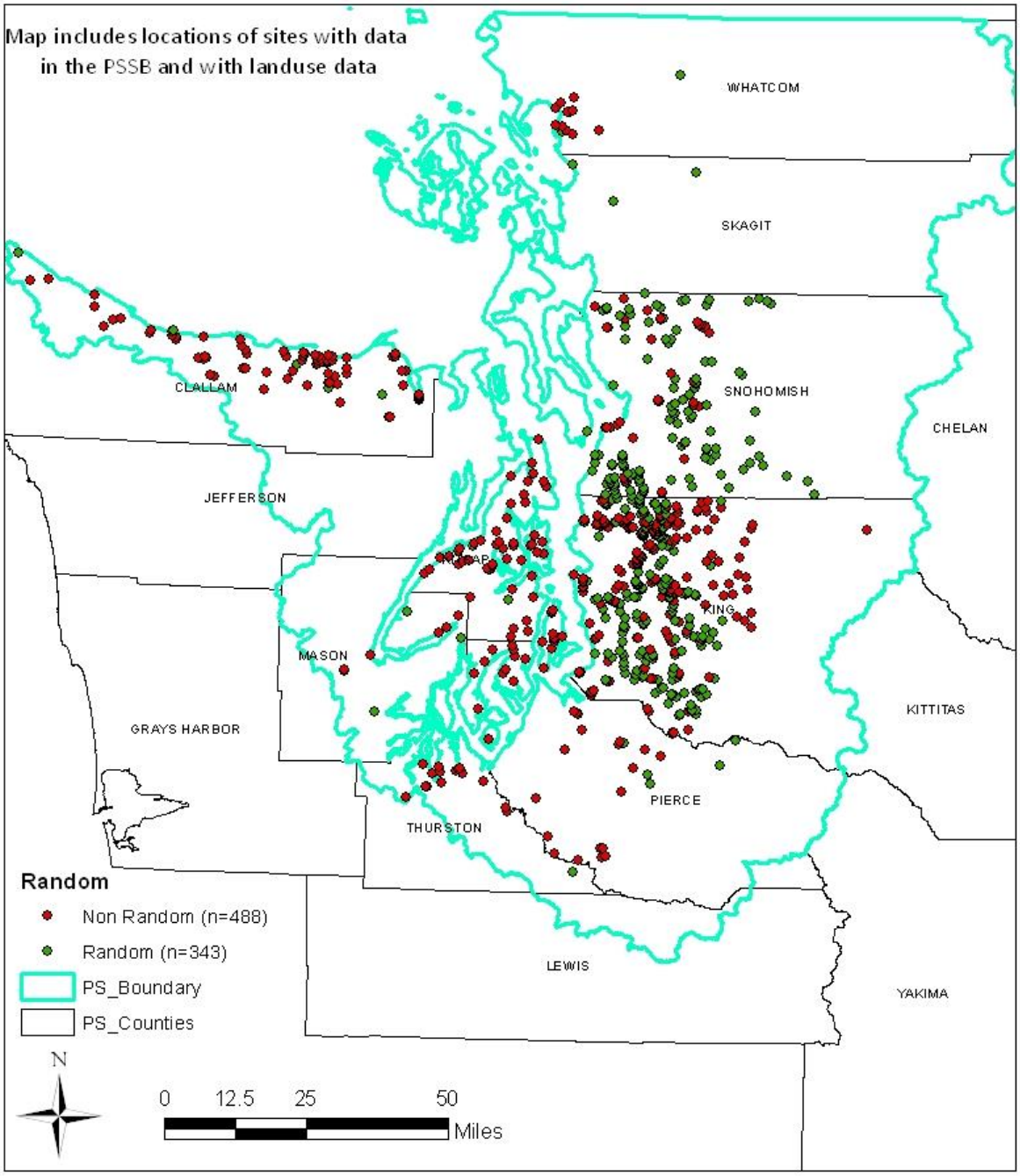


Figure 1. Distribution of sites from random sampling designs (random) versus other site selection methods (non random).

Land use/cover (LULC) for each county within Puget Sound, and Puget Sound in its entirety (far right bar) are summarized in Figure 2 with the number of random and non-random sites within each County overlaid. Percent forest is the dominant LULC throughout Puget Sound. King and Snohomish Counties have the most sites, but if all sites are considered, Clallam, Kitsap, Pierce, Thurston, and Whatcom counties each have at least 16 sites with BIBI data available.

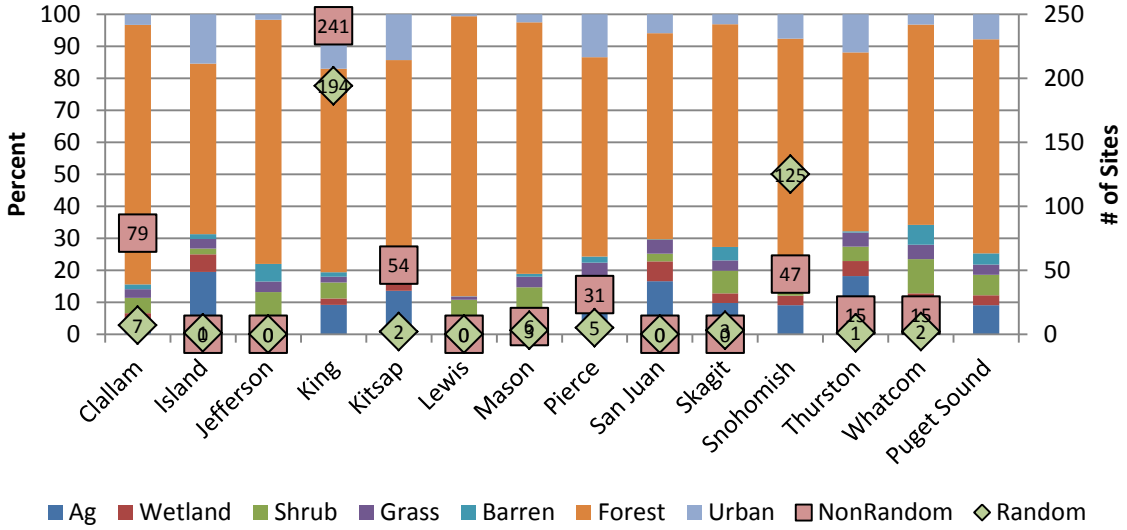


Figure 2. LULC for the portion of each County located within Puget Sound and the number of random and non-random sites with BIBI data stored in the PSSB.

LULC for each WRIA within Puget Sound are summarized in Figure 3, also with the number of random and non-random sites overlaid. WRIA 8 (Cedar-Sammamish) has the most number of sites.

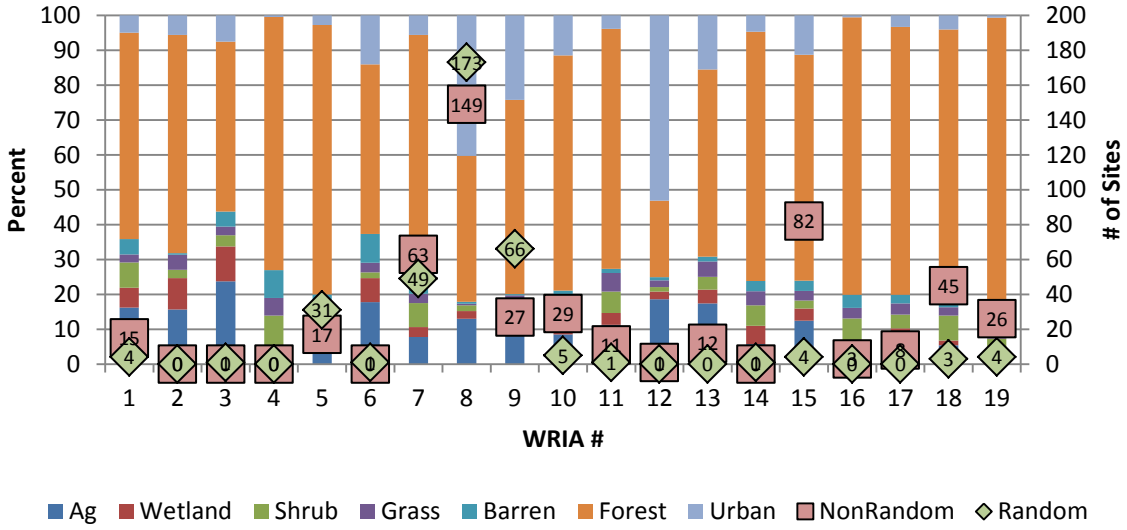


Figure 3. LULC for the portion of each WRIA located within Puget Sound and the number of random and non-random sites with BIBI data stored in the PSSB. WRIA 1 – Nooksack, 2 – San Juan, 3 – L. Skagit/Samish, 4 – U. Skagit, 5 – Stillaguamish, 6 – Island, 7 – Snohomish, 8 – Cedar/Sammamish, 9 – Duwamish/Green, 10 – Puyallup/White, 11 – Nisqually, 12 – Chambers/Clover, 13 – Deschutes, 14 – Kennedy/Goldsborough, 15 – Kitsap, 16 – Skokomish/Dosewallips, 17 – Quilcene-Snow, 18 – Elwha/Dungeness, 19 – Ivre/Hoko.

Both the histogram and the cumulative frequency distribution of % urban land use in the watershed for random, non-random, and all sites combined (All) have similar shapes to each other (Figure 4). All three data sets are over weighted in sites with less than 10% urban land use within the watershed and underweighted in sites with greater than 90% urban land use.

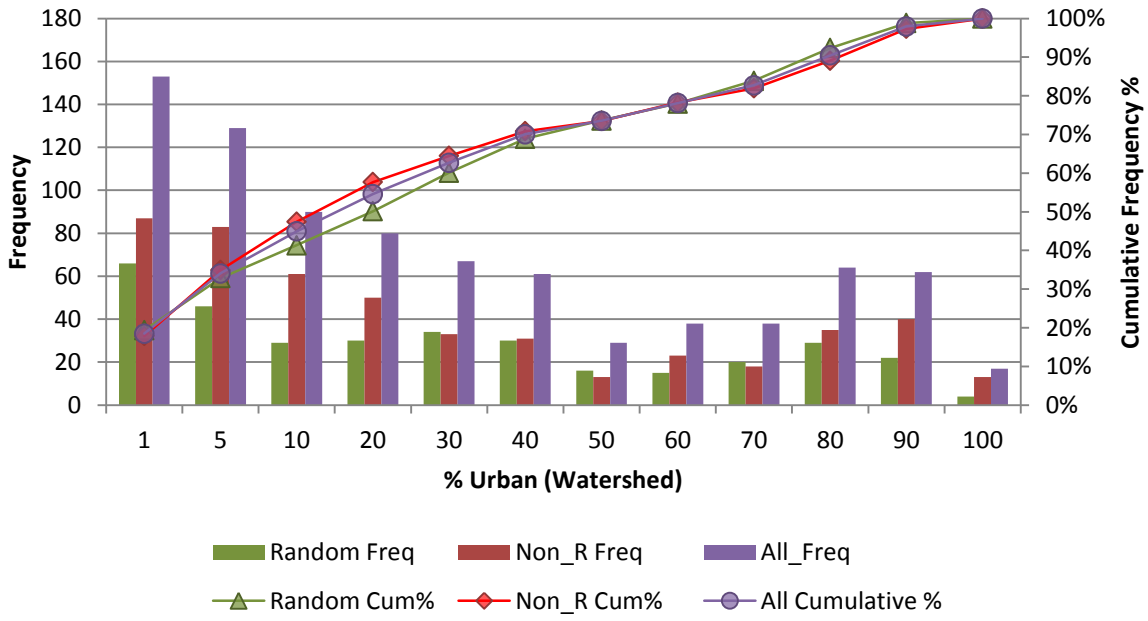


Figure 4. Histogram of random (blue), nonrandom (red) sites, and all sites combined (purple). The bins range from 0-1, 1-5, 5-10, and then switch to 10 % intervals (10-20, 20-30, etc.). The number shown on the x-axis represents the upper end of the bin range.

Regressing the cumulative frequency distribution of the random sites versus either the non-random or all sites combined yields a very tight fit (Figure 5).

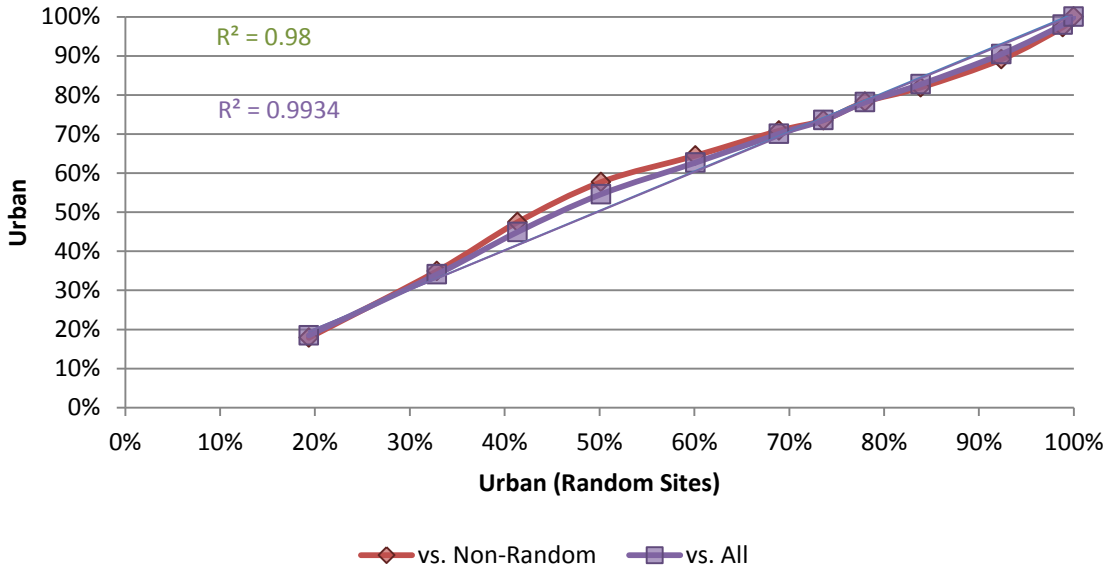


Figure 5. The cumulative frequencies from the random sites were graphed against the cumulative frequencies of the non-random (red) and all sites combined (purple). The R^2 values listed are for the best fit line with y-intercept of 0.

Another potential consideration is the total surface area sampled (Figure 6, Table 1). The 3 ft^2 sites are limited almost exclusively to Snohomish, King, and Pierce Counties.

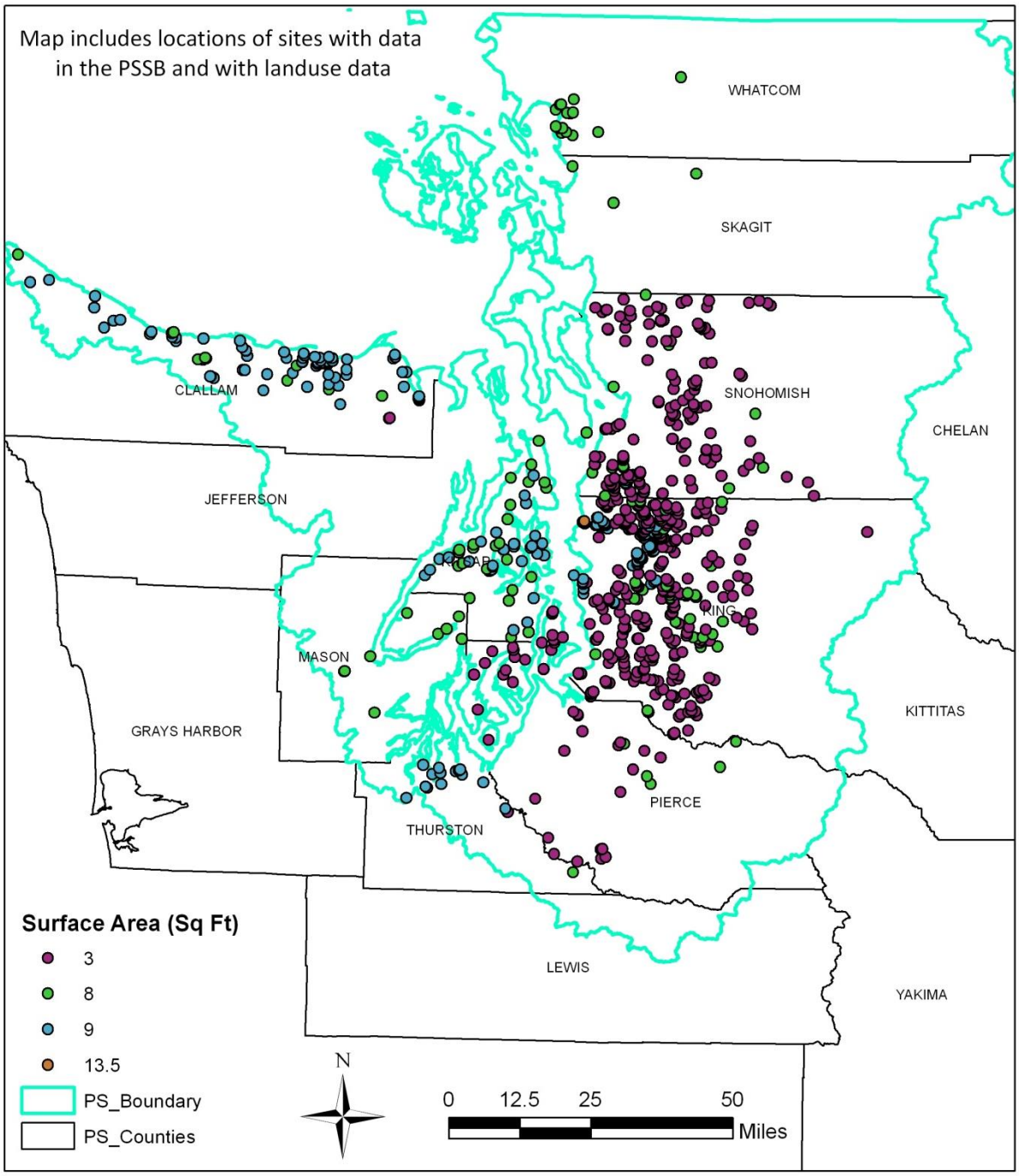


Figure 6. The distribution of sites by total surface area sampled.

Table 1. The number of sites for each total surface area sampled based on random and non-random sampling designs.

Surface Area (ft^2)	Non-Random	Random	Total
3	231	254	485
8	56	87	143
9	196		196
13.5	4		4
Total	**487**	**341**	**828**

Figure 7 attempts to combine the information regarding surface area and site selection design into one map.

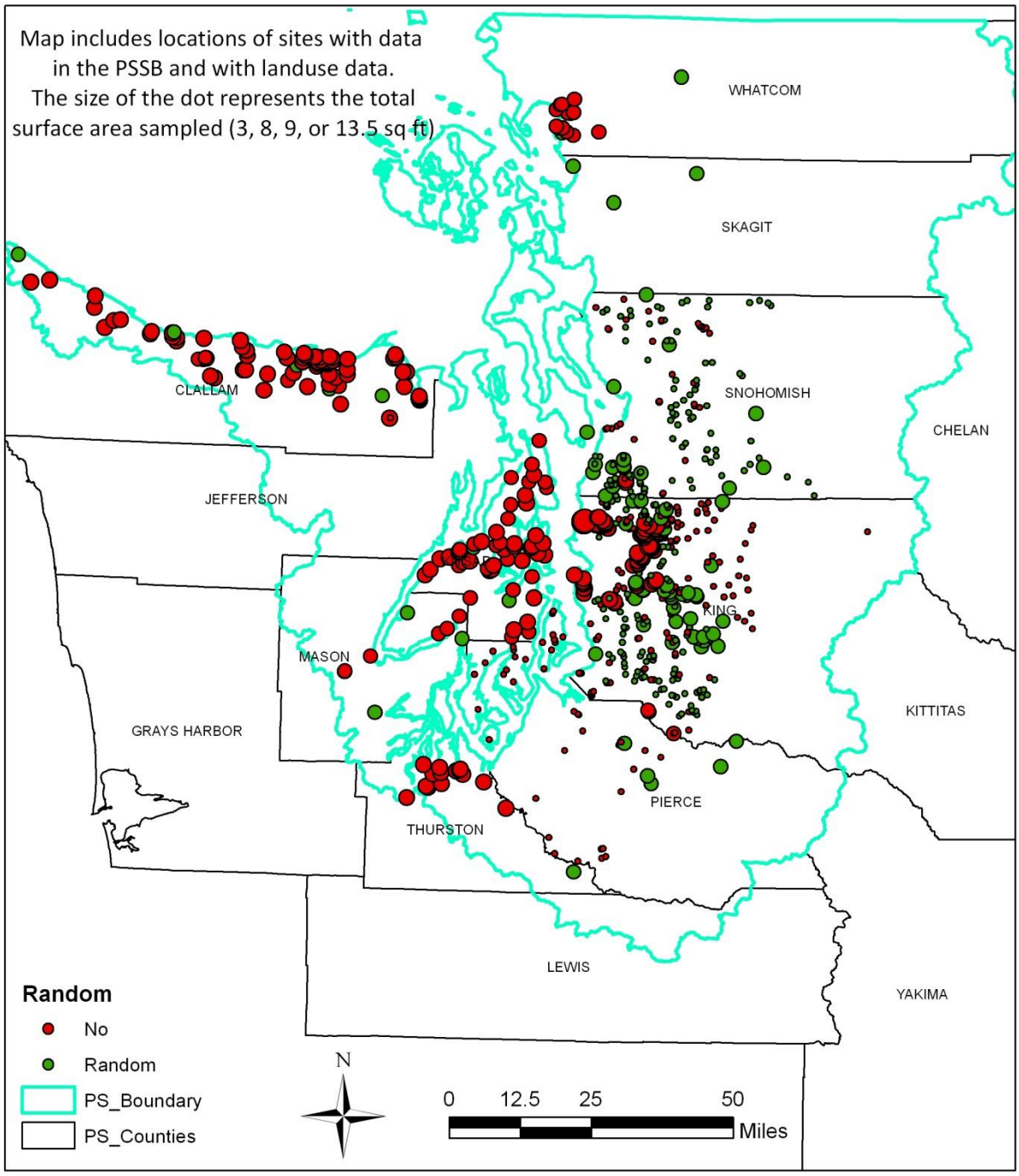


Figure 7. The distribution of sites by random vs. non-random with the size of the dot representing the total surface area sampled (3, 8, 9, or 13.5 ft^2).

The histograms based on site selection method and surface area are difficult to interpret, however it does seem that the nonrandom, 8 sq feet sites generally have less urban in the watershed (Figure 8). The cumulative frequency diagrams suggest that most of the site distributions are fairly consistent regardless of site selection method or surface area sampled (Figure 9). However, there are two exceptions: the nonrandom, 8 ft^2 sites and the nonrandom, 9 ft^2 sites. The nonrandom 8 ft^2 sites have a

higher proportion of sites with 10-40% urban and very few sites with ≥ 50% urban compared to the other categories. In contrast, the non random 9 ft^2 sites have a larger proportion of sites in areas with >70% urban than the other categories.

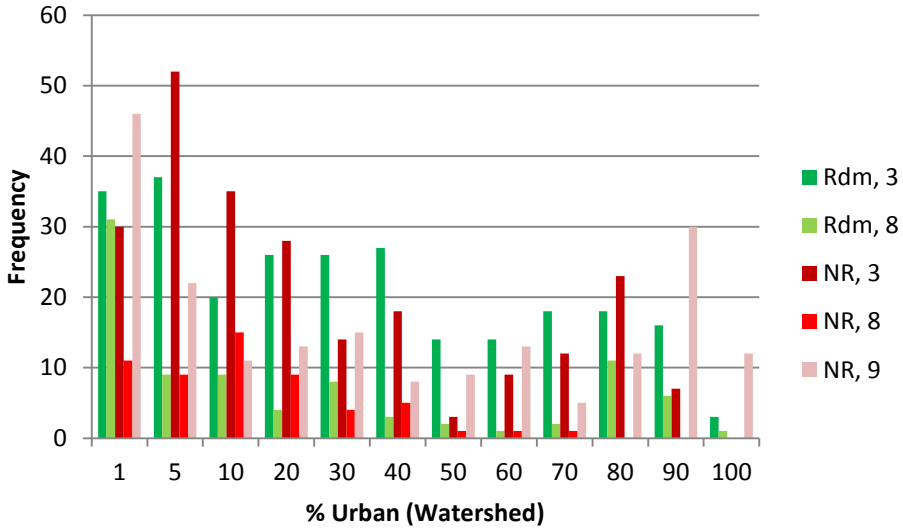


Figure 8. Histogram of % urban in the watershed for random (rdm) and non-random (NR) site selection in addition to total surface area sampled (3, 8, and 9 sq feet).

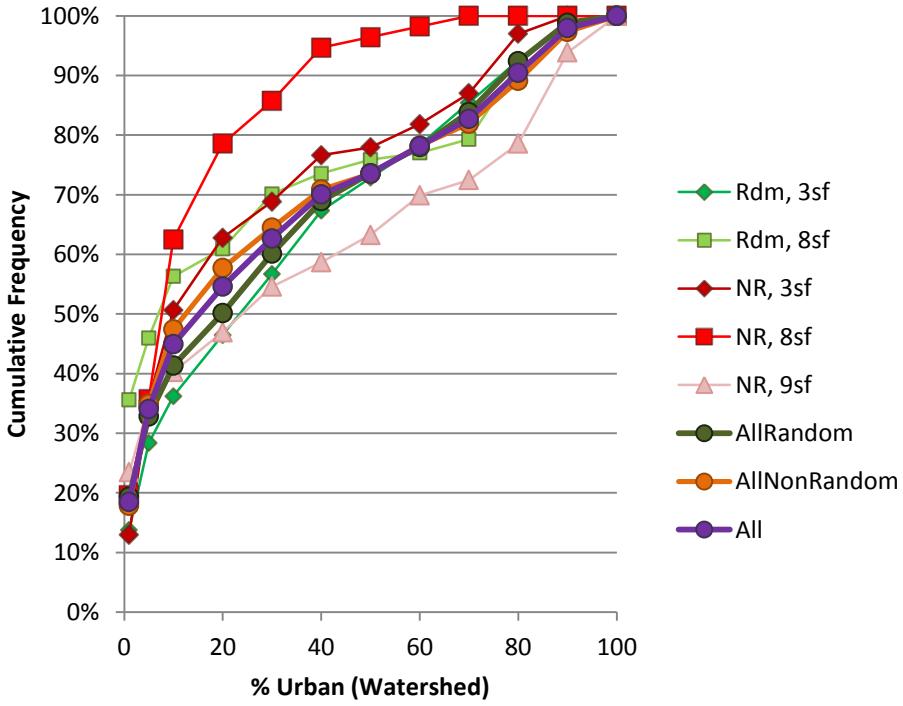


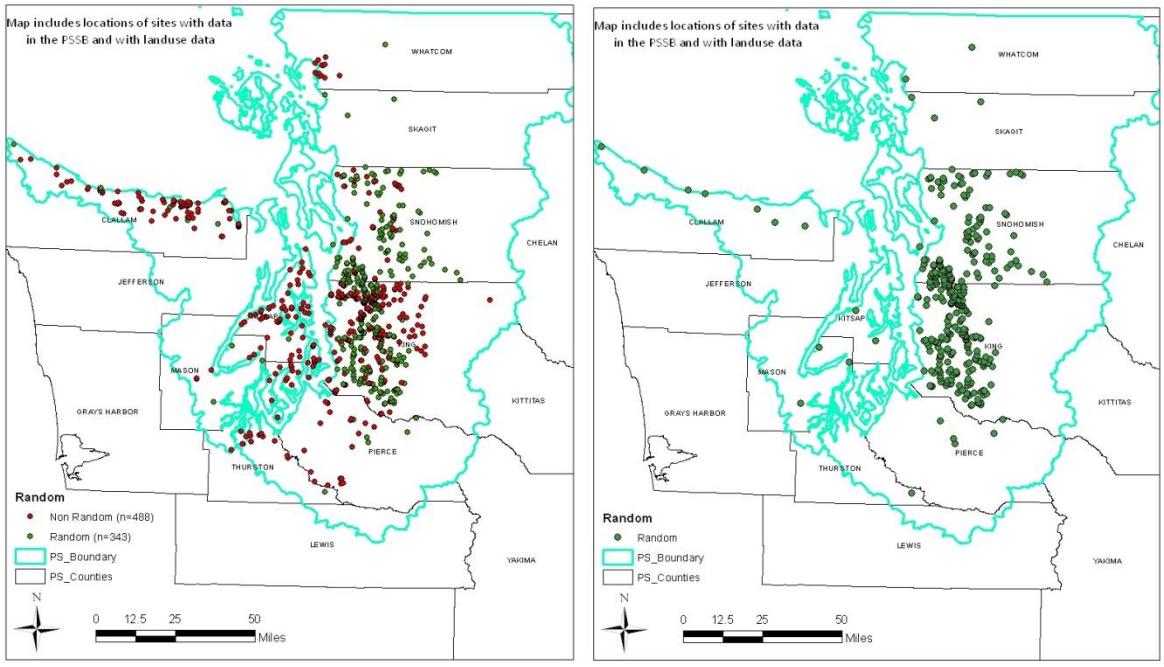
Figure 9. Cumulative frequency diagram of % urban in the watershed for random (rdm) and non-random (NR) site selection in addition to total surface area sampled (3, 8, and 9 sq feet). All sites combined are also graphed (purple).

Table 2 summarizes the distribution of sites by jurisdiction.

Table 2. Breakdown of site selection method (random vs. non-random) and total surface area sampled (3, 8, 9, 13.5 ft^2) by jurisdiction.

Jurisdiction	Non-Random				Random		Total
	3	8	9	13.5	3	8	
Adopt-A-Stream Foundation	7						7
City of Bainbridge Island			6				6
City of Bellevue			17				17
City of Bellingham		14					14
City of Everett	6						6
City of Federal Way	11						11
City of Issaquah	9						9
City of Kirkland	7						7
City of Lake Forest Park	4						4
City of Redmond			33				33
City of Seattle			31	4			35
Clallam County	1		72				73
King County - DNRP	39				157	59	255
King County - Roads	82						82
Kitsap County		25	21				46
Pierce County	33						33
Skokomish Tribal Nation		3					3
Snohomish County	32		2		97		131
Thurston County			14				14
Washington State Department of Ecology		14				28	42
Grand Total	**231**	**56**	**196**	**4**	**254**	**87**	**828**

Other Info for consideration:



Side by side comparison of all sites (left) and random only (right)

Site Selection – Follow up

Jo Wilhelm, King County, 4/5/2012

Cumulative Frequency Diagrams

First I added the 28 Ecology sites (Random, 8 ft^2). The Ecology data set is weighted heavily towards sites with very little urbanization within the contributing basin and has a unique land use distribution compared to the other data sets (Figure 1):

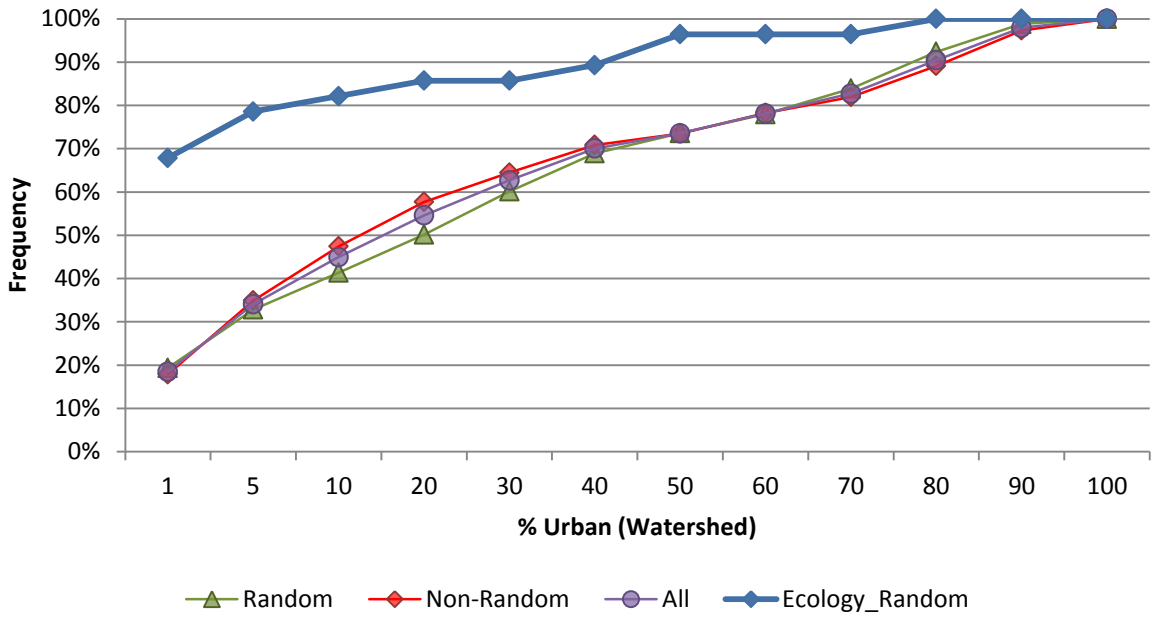


Figure 1. Cumulative frequency diagram of percent urban within the contributing watershed for random (n=487), non-random (341), all (828), and Ecology's random sites (28).

Secondly, I added all possible permutations onto the cumulative frequency figure. Good luck making sense of this (Figure 2)! There was a method to my madness in terms of naming and color coding. The first name refers to site selection (random, nonrandom, or all); the second name refers to the surface area (3, 8, 9, 8+9 ft^2, or all). Randoms are all graphed in shades of green; non randoms in shades of red; the combined random and non-random in shades of purple/blue. Of note: combining the 8 and 9 ft^2 sites does put those lines (both for non-random and all; there were no random 9 sites) in the mix with the 'typical' cumulative frequency distribution. **If I send this document as a word file, I am noticing that the graph is a living creature. You should be able resize the figure and also click on a data set and delete what you don't want to see. Therefore, feel free to play with what you see below to pick out the most relevant information.**

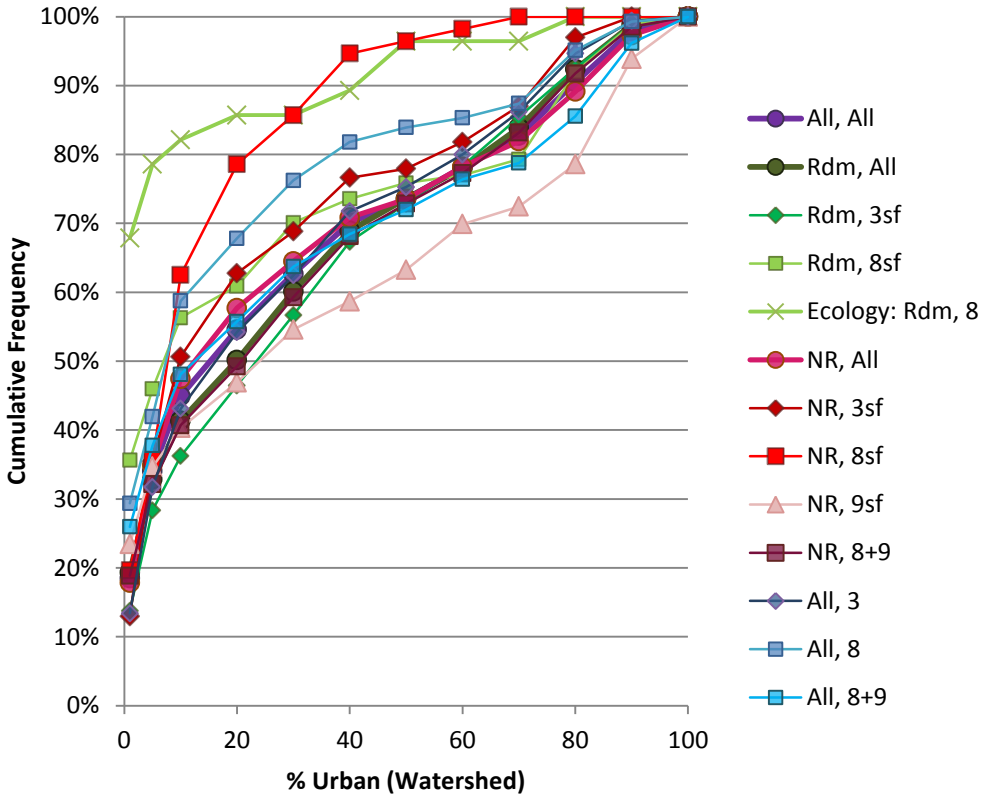


Figure 2. Cumulative frequency diagram of percent urban within the contributing watershed for varying combinations of site selection (random, nonrandom, all) and surface area (3, 8, 9, 8+9 ft^2, and all).

Thirdly, here is my reduced chart (Figure 3). I would recommend going forward with recalibrating the BIBIs on all the 3 ft^2 data (blue diamond) and all the 8 plus 9 ft^2 data (blue square) regardless of site selection method. If we want to control the distribution of landuse used for recalibration, we could just select a subset of the data by randomly drawing a specified # of sites from each 5% urban bin.

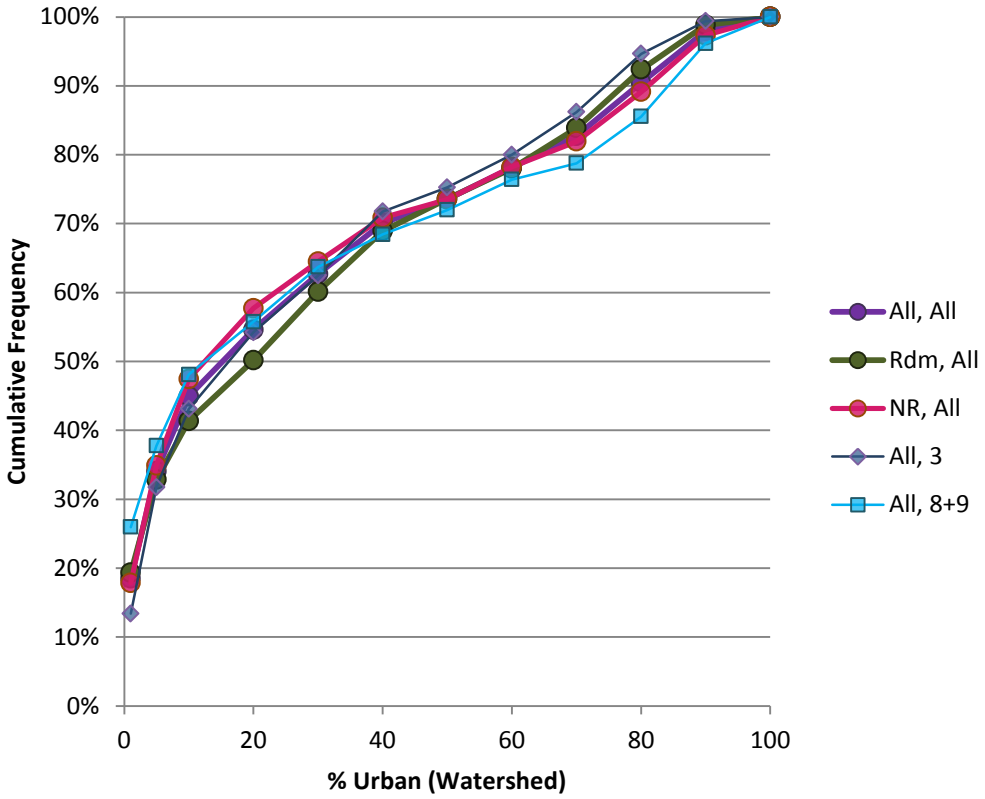


Figure 3. Cumulative frequency diagram of percent urban within the contributing watershed for varying combinations of site selection (random, nonrandom, all) and surface area (3, 8, 9, 8+9 ft^2, and all). My proposal would be to recalibrate the BIBI based on All, 3 for the 3 ft^2 methods and All, 8+9 for the 8 ft^2 methods. Both of these cumulative frequency distributions follow a similar distribution to all the data sets together.

Elevation

There are several sites with BIBI data that are above 500 m (Figure 4). I don't have an elevation shape file for all of Puget Sound, so I don't have a quick way to do a landuse analysis for the Puget Sound basin with elevation < 500m. I'll see if someone in-house can help, but if not we'll have to consider whether we want to see if Peter is able and willing to help.

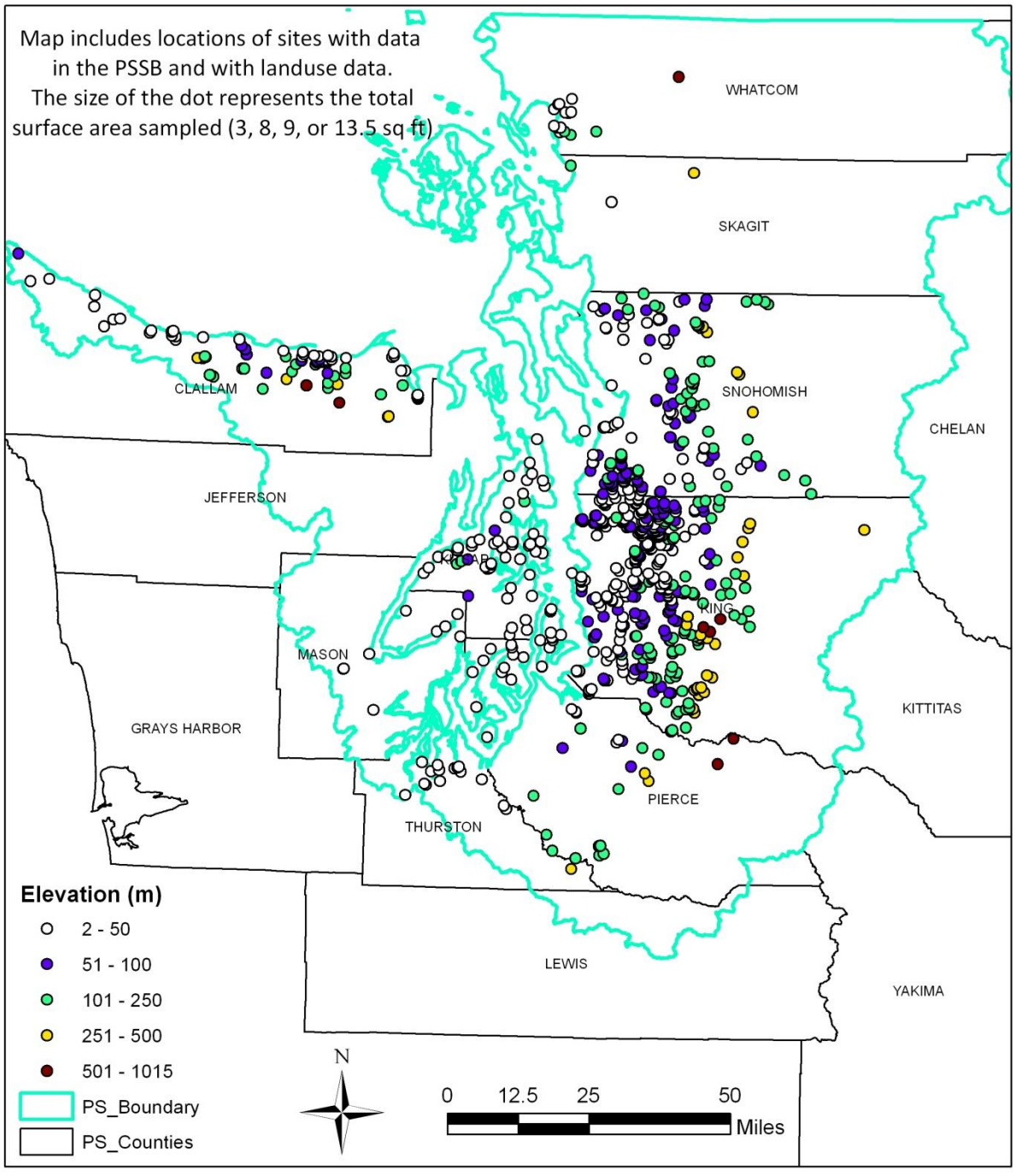


Figure 4. Elevation in meters of sites with BIBI and landuse data.