Supplementary material
Supplementary Materials for:
Landscape Scale Variation in the Hydrologic Niche of California Coast Redwood

This file includes:
Appendix 1 Supplementary Text
Supplementary Methods:
Logging History of the Three Sites
Interpolated Height above a Stream (IHAS)
Soil Water Storage
Relative Influence Metric of Boosted Regression Trees (BRT)
Influence of the Interaction between Fog and Belowground Moisture on Redwood Distributions
Diurnal Variation in Fog Cover
Analyses of the Variation in Model Performance Across Scales and Variability within Coarser Grain Sizes

Supplementary Results:
Cross-Correlations of Soil Water Storage with Other Variables
Interaction of Fog and IHAS

Supplementary Discussion: Differences in the Importance of the Predictor Variables Across the Sites: Site-Specific Idiosyncrasies and Possible Latitudinal Trends
Soil Water Storage: Ecologically Important at Big Basin, Confounding Variable at Mt. Tam
Differences in the Range and Importance of Fog Frequency Across the Three Sites Interactions of Fog and IHAS Predict Redwood Distributions

Appendix 2 Supplementary Figures
Figure A1: Redwood probability as a function of the interaction of fog frequency and interpolated height above a stream (IHAS).

Figure A2: Cross-correlations of topographically-based measures of hydrology and fog from Mt. Tam.

Figure A3: Cross-correlations of topographically-based measures of hydrology and fog from Big Basin.

Figure A4: Cross-correlations of topographically-based measures of hydrology and fog from Jackson Forest.

Figure A5: Relationship of soil water storage and fog frequency at Mt. Tam.
Figure A6: Relationship of soil water storage and IHAS at Big Basin.

Figure A7: Relationship of redwood density and nighttime fog hours.

Figure A8: Maps of fog frequency and redwood distributions close to the coast at Mt. Tam.

Figure A9: Maps of fog frequency and redwood distributions close to the coast at Big Basin.

Figure A10: Maps of fog frequency and redwood distributions close to the coast at Jackson Forest.

Figure A11: Performance of species distribution models for redwoods developed at 6 different resolutions.

Figure A12: Performance of species distribution models for redwoods developed with data at 6 different resolutions and tested on a 10-m resolution distribution dataset.

Figure A13: The relative influence of each different topographic predictor for models of redwood distributions developed at different grain sizes.

Figure A14: Predicted probability of redwood occurrence from models developed at varying grain sizes.

Table A1: Performance of a machine learning model predicting the presence and absence of redwoods from five metrics of belowground moisture and fog at three sites.

Figure A15: Variability of predictor variables at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at Mt. Tam.

Figure A16: Variability of predictor variables at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at Big Basin.

Figure A17: Variability of predictor variables at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at Big Basin.

Figure A18: Variability of model output probabilities at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at all three sites.
Appendix 1 Supplementary Text

Logging History of the Three Sites
Jackson Forest is a site of active logging, with stand age varying from ten to one hundred years old within the landscape, but also has one old-growth stand that has never been harvested. Mt. Tam contains Muir Woods National Park, an old-growth redwood forest. The remainder of the site covers forests that were logged in the early 1900’s. Big Basin harbors a small section of old-growth forest, and the remainder of the forest ranges from thirty to two hundred years old, depending on the date of the most recent harvest (Francis, 2013).

Interpolated Height above a Stream (IHAS)
Under the assumption that surface water pools where the elevation of the water table is above the elevation of the land surface (Dunne and Black, 1970), IHAS is effectively the depth to the water table for a given location on a topographic surface. However, IHAS is not the depth at which the roots would need to grow to access water. Due to hillslope processes, IHAS is likely to be negatively correlated with soil moisture in unsaturated soil. Higher soil moisture in lower IHAS locations results from physical soil erosion along hillslopes causing deeper soils at lower IHAS, chemical soil erosion along hillslopes causing soils with higher water storage capacity at lower IHAS (Gessler, Chadwick, Chamran, Althouse, and Holmes, 2000; Jenny, 1941), and drainage causing higher water input to soil at lower IHAS (Beven and Kirkby, 1979). Negative IHAS values occur when the ground elevation is below the modeled stream elevation. They are areas where soil moisture would be expected to accumulate and can be interpreted as an extension of a linear IHAS scale. 1.5% of the total values of IHAS at Mt. Tam, 2.2% of the total IHAS values at Jackson Forest, and 2.7% of the total values of IHAS at Big Basin were negative values.

Soil Water Storage
Variation in soil water storage capacity results from both topo-edaphic variation as discussed above, as well as underlying geologic variation. We expected geologic variation to result in variation in soil type that was independent of topography.

Relative Influence Metric of Boosted Regression Trees (BRT)
The relative-influence metric indicates how many times a variable is selected for splitting, weighted by the squared improvement to the model that results from the split, averaged over all of the trees.

Influence of the Interaction between Fog and Belowground Moisture on Redwood Distributions
We visualized interactions between two variables by splitting the data into categories corresponding to each of 8 evenly-spaced bins of fog frequency and 8 evenly-spaced bins of IHAS, then calculated the mean redwood probability within each category. 8 bins rather than deciles were used for the multi-variable analysis to maintain a high sample size for each multivariate bin. For presentation, the minimum value of the first IHAS bin was set to zero, therefore excluding the negative values of IHAS. One multivariate bin at Big Basin had only six pixels with which to estimate the mean and was excluded (all others had greater than 200
Diurnal Variation in Fog Cover
Fog frequency varies throughout the day, as warm temperatures cause fog to evaporate. To make sure that diurnal variation in fog cover did not affect the results, we also conducted a subset of the analysis with a 4-km map of daytime and nighttime fog hours (Torregrosa et al. 2016). We calculated the plotting the mean and bootstrapped 95% percent confidence intervals of redwood density within each twentieth percentile of nighttime fog hours at each site.

Analyses of the Variation in Model Performance Across Scales and Variability within Coarser Grain Sizes
To assess how the performance of the models differed across grain sizes, we rescaled the predictor and the response variables to different grain sizes and re-calculated the AUC, overall accuracy, specificity, sensitivity, and relative influences. All of the predictor variables were re-scaled by calculating the mean of all of the 10 m x 10 m resolution cells within each larger resolution cell (0,50,100,150,200,250). To calculate the interquartile ranges of each predictor variable, the first quartile and third quartile of all of the 10 m x 10 m pixels within an 800 m x 800 m grain size was calculated. To calculate the interquartile range, for each 800 m x 800 m pixel, the difference between the third and first quartile was calculated. To calculate the range of predicted probability of redwood occurrence within each 800 m x 800 m cell, the minimum and the maximum predicted probability of occurrence at 10 m x 10 m within each 800 m x 800 m cell was calculated. For each 800 m x 800 m grid cell, the minimum was subtracted from the maximum to obtain the range. We then plotted the distribution of the ranges across all of the 800 m x 800 m cells. All rescaling of variables, and calculation of the minimum and maximum, and first quartile and third quartiles, was implemented in GDAL (GDAL/OGR 2020).

Supplementary Results
Cross-Correlations of Soil Water Storage with Other Variables
At Big Basin, soil water storage was modestly correlated with IHAS (Fig. A3), and soil water storage declined with increasing IHAS (Fig. A6). Soil water storage and fog frequency were modestly correlated at Mt. Tam (Fig. A2), and increased from fog frequency of approximately 0.10 to approximately 0.55 (Fig. A5).

Interaction of Fog and IHAS
In sites with low fog frequency, redwood probability was highest at the lowest IHAS positions (Fig. A1). In sites with moderate fog frequency, redwood probability was highest at the second-to-lowest IHAS position. Jackson Forest, the northernmost site, had the highest mean annual precipitation from rainfall at 1,214 mm and the lowest mean annual temperature at 13.14 °C (Table 1). In microsites at Jackson Forest with 8-25% fog frequency, redwood probability was highest at 0-40 m IHAS, whereas above 25% fog frequency, redwood probability was higher at 40-80 m rather than 0-40 m IHAS (Fig. A1). At Mt. Tam, fog frequency was highest, ranging from approximately 5-55%, mean annual precipitation from rainfall was slightly lower than Jackson Forest at 1,160 mm, and mean annual temperature was slightly higher than
Jackson Forest at 13.96°C (Table 1). At Mt. Tam, in microsites below approximately 27% fog frequency, redwood density was highest at 0-40 m IHAS whereas between approximately 27 and 33% fog frequency, redwoods were most abundant at 40-80 m IHAS (Fig. A1). At Big Basin, fog frequency was the lowest of all the sites (Fig. A1 and 1). Mean annual precipitation from rainfall was also lowest at 1,093 mm, and mean annual temperature was highest at 14.72 °C (Table 1). At Big Basin, redwood probability was highest at 0-40 m IHAS in all bins of fog frequency (Fig. A1).

Supplementary Discussion: Differences in the Importance of the Predictor Variables Across the Sites: Site-Specific Idiosyncrasies and Possible Latitudinal Trends

The three sites in our study cover a latitudinal gradient approximately one-third of the total latitudinal range of coastal redwood. This latitudinal gradient corresponds to a gradient in temperature and precipitation – Jackson Forest had the lowest temperature and the highest precipitation, and Big Basin had the highest temperature and the lowest precipitation. We expect that some of the differences across the sites may be a result of site-specific idiosyncrasies, while other differences across the sites may be reflective of latitudinal trends, as discussed below.

Role of Soil Water Storage Capacity in Redwood Distributions: Ecologically Important at Big Basin, Result of a Correlation with another Variable at Mt. Tam

The response of redwood density to soil water storage varied across the three sites. Soil water storage was the most important variable in the BRT model at Big Basin, and the second-most important variable in the BRT model at Mt. Tam (Fig. 3). However, examination of the functional relationships of redwood density to soil water storage suggests that the relatively high influence of soil water storage at Mt. Tam is an artifact or the result of a correlation with another variable that is important to redwood distributions. At Mt. Tam, redwood probability declined at soil water storage capacity above 8 cm (Fig. 3). We suspect this may be due to a correlation with other factors that may have been limiting to redwood distributions. Soil water storage capacity above 10 cm was positively correlated with fog frequency (Fig. A2), therefore, the sites with 10-15 cm water storage, which also had very low redwood densities (Fig. 4, bottom middle plot) sites with high soil water storage capacity also had uniquely high fog frequencies (Fig. A5). These sites were close to the coast (Fig. A8), and we hypothesize that they were inhospitable to redwoods because of high fog frequency, fast winds, and high soil salt content.

On the other hand, we believe that soil water storage is an ecologically important factor in determining redwood distributions at Big Basin. The high relative influence of soil water storage at Big Basin (Fig. 3) is consistent with the steep increase in redwood density with increasing soil water storage at Big Basin (Fig. 4). Big Basin also had the lowest soil water storage values of any of the sites, corresponding to very low redwood densities (Fig. 4). The low soil water storage values at Big Basin is consistent with knowledge of large deposits of sandy soils at Big Basin, which favor xeric plant communities (Griffin, 1964).
Differences in the Range and Importance of Fog Frequency Across the Three Sites

In the case of fog frequency, we expect that differences in the ranges of fog frequency across the site may be attributable to a latitudinal gradient, which may partially explain why fog frequency was less important to redwood distributions at Big Basin in comparison with the other two sites. Fog frequency was much lower at Big Basin than at the other two sites (Fig. 4). Fog frequency varies throughout the day, and tends to be higher in the evenings and mornings, until warm temperatures cause fog to evaporate during day. Our fog frequency map collects data at 10:30 AM, which is known to be close to the time of the day when fog ‘burns off’ (Fischer et al. 2009). Therefore, by 10:30 AM in Big Basin, fog may have already evaporated, whereas at the other two sites, it may remain cool enough for fog to remain further into the morning. To test this hypothesis, we re-made Figure 4 using a fog frequency map that shows the hours of nighttime fog (Figure A7). This map is available at 4 km x 4 km resolution, and given the small size of our landscapes, there were fewer 4 km x 4 km grid cells to analyze (particularly for Mt. Tam, the smallest site in our study). However, Big Basin shows a similar range in nighttime fog hours as Jackson Forest (in Figure A7, compare x-axis range in the leftmost plot in Figure A7 to x-axis range in the rightmost plot). This confirms our hypothesis that redwoods at Big Basin are exposed to a gradient in fog cover, even if not detectable the fog frequency map. However, redwoods at Big Basin still show a far less convincing increase in density from low to high fog than do the other two sites, regardless of which fog map is used (Fig. A7 and Figure 4). As a final explanation, it is possible that fog frequency has declined significantly at Big Basin. If the redwoods at Big Basin germinate and matured under a regime where fog frequency was consistently high across the site, then it is plausible that their distributions would be primarily limited by low soil water storage and height above a stream, rather than fog. We conclude that the variability in the importance of the spatial patterns of fog frequency to redwood distributions along a latitudinal gradient is a topic in need of further study.

Interactions of Fog and IHAS Predict Redwood Distributions

The influence of the interactions of IHAS and fog frequency on redwood abundance suggest that redwoods may be able to substitute fog for soil moisture along spatial gradients in soil moisture availability and fog cover. This interpretation builds upon findings from stable isotope analysis that redwood can substitute fog for water from rainfall during years when rainfall is low (Dawson, 1998). At the warmest, driest, and least fog-influenced landscape, redwoods were always most abundant in plots with the lowest IHAS (Table 1, Fig. 4). Conversely, at Jackson Forest and Mt. Tam, redwood abundance remained high at higher IHAS positions where fog was moderate, and redwoods were more abundant at higher IHAS positions in sites with high fog frequency at Jackson Forest (Fig. 4). Jackson Forest and Mt. Tam had higher rainfall (Table 1) and higher fog frequency, and Mt. Tam can be water-logged for some portions of the year (Carroll, 2014). It is possible that the lower occurrence of redwoods in sites at low IHAS and sufficient fog frequency at Jackson Forest and Mt. Tam could indicate that sites with high water availability from both fog and streams could represent the wetter limit of the redwood hydrologic niche. As an alternative explanation, the elevation at which redwood densities are maximized at both of these sites could be the elevations at which fog drip is maximized, and could therefore have more water available than sites close to streams.
Appendix 2 Supplementary Figures

Figure A1: Redwood probability as a function of the interaction of fog frequency and interpolated height above a stream (IHAS). Heat maps show the probability of a redwood as the color of each box, calculated as the mean probability of redwood presence within each combination of percent fog cover (y-axis) and IHAS (x-axis). Areas with no color are locations where there is no combination of the two variables in the plot. All estimates had more than 200 pixels with which to estimate the mean value. Note the difference in the y-axis ranges across the three plots.

Figure A2: Cross-correlations of topographically-based measures of hydrology and fog from Mt. Tam. Scatterplot matrix shows the absolute value of the Pearson R correlation coefficient as the color between the variable labeled on the x-axis and the variable labeled on the top y-axis.
Figure A3: Cross-correlations of topographically-based measures of hydrology and fog from Big Basin. Scatterplot matrix shows the absolute value of the Pearson R correlation coefficient as the color between the variable labeled on the x-axis and the variable labeled on the top y-axis.

Figure A4: Cross-correlations of topographically-based measures of hydrology and fog from Jackson Forest. Scatterplot matrix shows the absolute value of the Pearson R correlation coefficient as the color between the variable labeled on the x-axis and the variable labeled on the top y-axis.
Figure A5: Relationship of soil water storage and fog frequency at Mt. Tam. The dark blue line shows the average soil water storage capacity within each tenth percentile of fog frequency at Mt. Tam. The dark blue shaded regions show the bootstrapped 95% percent confidence intervals of soil available water storage capacity calculated within each tenth percentile of fog frequency.
Figure A6: Relationship of soil water storage and IHAS at Big Basin. The dark blue line shows the average soil available water storage (cm) within each tenth percentile of IHAS at Big Basin. The blue shaded regions show the bootstrapped 95% percent confidence intervals of soil water storage calculated within each tenth percentile of IHAS.

![Graph showing relationship between soil water storage and IHAS.

Figure A7: Relationship of redwood density and nighttime fog hours. The dark blue line shows redwood density within each 20th percentile of hours of nighttime fog at each site. The dark blue line shows the mean and the blue shaded regions show the bootstrapped 95% percent confidence intervals of redwood probability calculated within each 20th percentile of GEOS-derived hours of nighttime fog. From Torregrosa et al. (2014).
Figure A8: Maps of fog frequency and redwood distributions close to the coast at Mt. Tam. Note the difference in the scale bar between the zoomed-in map (left) and the full map (right).

Figure A9: Maps of fog frequency and redwood distributions close to the coast at Big Basin.
Figure A10: Maps of fog frequency and redwood distributions close to the coast at Jackson Forest.

Figure A11: Performance of species distribution models for redwoods developed at 6 different resolutions. The predictor variables were spatially aggregated to coarser resolutions by calculating the mean of all of the 10-m resolution pixels within the region covered by the coarser-resolution pixel.
Figure A12: Performance of species distribution models for redwoods developed with data at 6 different resolutions and tested on a 10-m resolution dataset. The predictor variables were spatially aggregated to coarser resolutions by calculating the mean of all of the 10-m resolution pixels within the region covered by the coarser-resolution pixel. Performance measures were calculated on a 10% testing sample of the data which was not included in model development.

Figure A13: The relative influence of each different topographic predictor in models of redwood distributions developed at different grain sizes. Models were built using predictor variables and that were spatially-aggregated to each grain size shown on the y-axis.
Figure A14: Predicted probability of redwood occurrence from models developed at varying grain sizes. The upper left plot shows the original redwood distributions data at 10 m resolution, and the other plots show the predicted probabilities of occurrence from models developed at a range of grain sizes from 10 – 250 m. The grain sizes are shown as the title of each map. Note that the areas with no data also include areas that were removed after being identified as lakes or urban areas.

Table A1: Performance of a machine learning model predicting the presence and absence of redwoods from five metrics of belowground moisture and fog at three sites. The overall accuracy and AUC calculated on a 30% testing dataset are shown for each of the three sites (This is the same as Table 2 in the main manuscript, but using a testing dataset that was 30% of the observations rather than 10%).

<table>
<thead>
<tr>
<th>Site</th>
<th>Overall Accuracy</th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Basin</td>
<td>0.65</td>
<td>0.71</td>
<td>0.79</td>
<td>0.52</td>
</tr>
<tr>
<td>Mt. Tam</td>
<td>0.76</td>
<td>0.83</td>
<td>0.81</td>
<td>0.71</td>
</tr>
<tr>
<td>Jackson Forest</td>
<td>0.63</td>
<td>0.68</td>
<td>0.72</td>
<td>0.55</td>
</tr>
</tbody>
</table>
Figure A15: Variability of predictor variables at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at Mt. Tam. Histograms show the interquartile range of all 10 m x 10 m pixels of each predictor variable within an 800 m x 800 m aggregation block area. Each bar on the histogram shows the number of 800 m x 800 m pixels within each bin of each interquartile range value shown on the x-axis.
Figure A16: Variability of predictor variables at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at Big Basin. Histograms show the interquartile range of all 10 m x 10 m pixels of each predictor variable within an 800 m x 800 m aggregation block area at Big Basin. Each bar on the histogram shows the number of 800 m x 800 m pixels within each bin of each interquartile range value shown on the x-axis.

Figure A17: Variability of predictor variables at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at Jackson Forest. Histograms show the interquartile range of all 10 m x 10 m pixels of each predictor variable within an 800 m x 800 m aggregation block area. Each bar on the histogram shows the number of 800 m x 800 m pixels within each bin of each interquartile range value shown on the x-axis.
Figure A18: Variability of model output probabilities at a fine-scale (10 m x 10 m) within a coarse grain size (800 m x 800 m) at all three sites. Histograms show the range of all 10 m x 10 m pixels of predicted redwood habitat suitability within an 800 m x 800 m aggregation block area. Each bar on the histogram shows the number of 800 m x 800 m pixels within each bin of each range value shown on the x-axis.

References


