Axel Malahieude
UID 204915219
7/31/2020

Finding the Best Vineyard Locations in California

View this project in video format: https://youtu.be/nLeyDbTNjc8

California has become increasingly well-known for its wine industry, which dominates US wine exports. From Napa Valley to the Southern California Wine Country, wine sales bring in over $43 billion to California, a number which has increased by 30% in the past decade alone (Horiuchi). However, California is an extremely geographically diverse state, so given the wine industry's economic impact, it is important to understand where wine can reliably be grown in California and whether current wineries are have optimized their production by selecting a suitable site. As such, the purpose of this project is to identify the best locations for grapevine growing and evaluate how well wineries have chosen their vineyard sites. This will also be useful for determining the potential for expansion and growth of the California wine industry.

1. Methodology

In order to find out the best places for grapevines, we need to know which criteria are important when selecting a vineyard site. Therefore, the first step to this project is to perform some research. To manage this project effectively, the two most important factors that determine grapevine production will be chosen. Afterwards, data will be collected for each of these growth factors, processed, and then prepared into a map to show their spatial characteristics throughout California. Then, based on the requirements of the growth factors, a unitless index will be created for each of those factors. These indices will allow for the aggregation of both growth factors into a single index to create a map that displays how suitable sections of California are for vineyards according to these factors. The last step to this will be to overlay current winery locations and see how well they have chosen their sites.

In summary, the steps involved are:

1. Research grapevine growth factors
2. Collect data related to the most important growth factors
3. Present the raw data in map format
4. Compute an index that measures grapevine suitability for each growth factor
5. Aggregate the indices into a single metric that describes overall grapevine suitability
6. Present this overall grapevine suitability data in map format
7. Analyze how current winery locations compare to these findings

2. Research Results

Preliminary research shows that there are a multitude of factors that affect grapevine production and quality, ranging between weather, geology, and geography. By cross-referencing multiple sources, the factors that have the greatest impact on grapevines are average temperature during the growth season and soil drainage levels (Johnson and Robinson), (Goldammer).

To evaluate temperature, researchers at UC Davis have created the heat summation index (Amerine and Winkler). This index is created by taking the average temperature for a region each day between April 1st and October 31st (the typical grape-growing season) and subtracting 50°F (the lowest temperature at which grapes can develop), then summing over all 214 days (CalWineries). The result is a value (called "degree days") that falls in the following scale:

Degree Days	Region	Grape Types
< 2500	I	Chardonnay, Pinot Noir, Gewurztraminer, Reisling
2500 - 3000	II	Cabernet Sauvignon, Merlot, Sauvignon Blanc
3000 - 3500	III	Zinfandel, Barbera, Gamay
3500 - 4000	IV	Malvasia, Thompson Seedless
> 4000	V	Thompson Seedless, other table grapes for eating

The general consensus is that as the number of degree days increases, the quality of wine decreases while the production capacity increases (Amerine and Winkler). In other words, there is a tradeoff between quality and quantity. Given that Region V mostly caters to table grapes and most well-known wines come from Region II, for the purposes of this project, we will take Region II to be the "best" temperature location. Next in the hierarchy will be Regions I and III, followed by Region IV and finally Region V.

As for soil drainage, the USDA has 7 classifications for drainage levels based on soil type. The 1st one is "very poorly drained" and the 7th one is "excessively drained." By reading the USDA Drainage Class Handbook and referencing viticulture resources, it was found that the optimal USDA drainage class for vineyards would be the 5th one, which is "well drained." For reference, more drainage means drier soil with less water availability, while less drainage means higher occurrence of waterlogged soil (Soil Survey Division Staff).

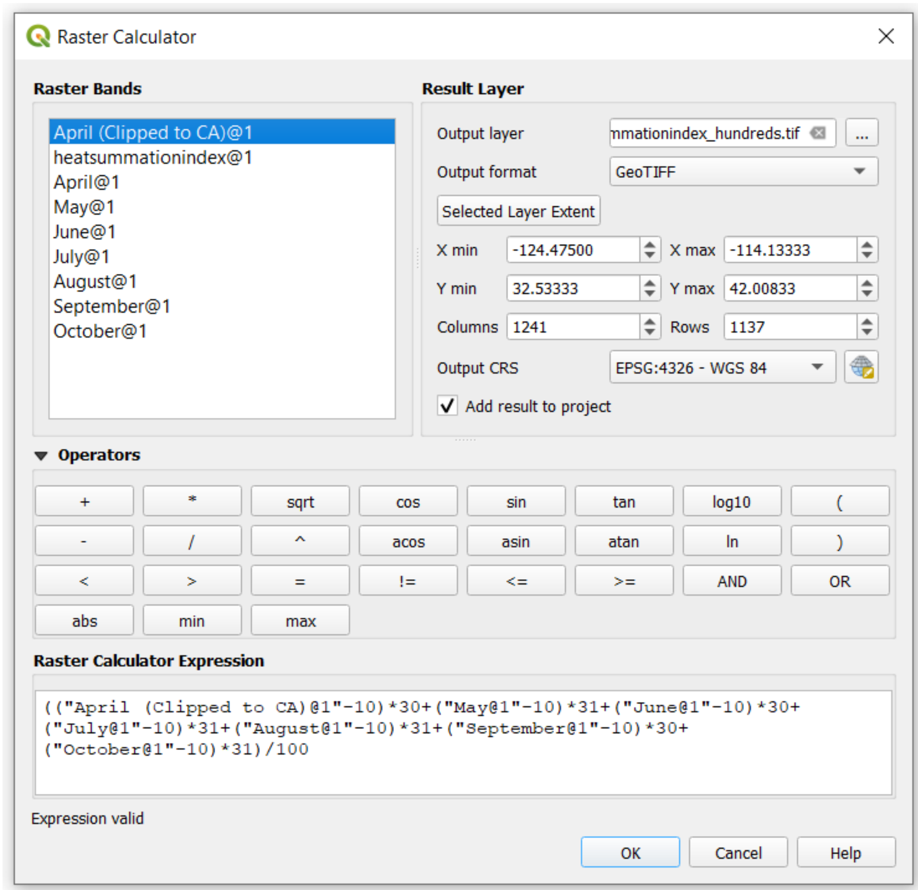
As a final remark, one may think that there are other more important factors affecting vine growth, such as rainfall or soil nutrient levels. Though such factors were historically important, they have since been overcome with modern agricultural techniques, such as by using fertilization and irrigation to get away with poor soil nutrients and lack of rain (Johnson and Robinson). As such, the focus of this project shifts to growth factors that cannot easily be modified, notably soil drainage and temperature.

3. Data Collection and Processing

3.1 Temperature Data

Historical average temperature data was collected from the WorldClim database in GeoTiff format at a resolution of 30 arcseconds (approximately 1 square kilometer). This data is per month, so 12 individual GeoTiff files were included. Unfortunately, historical daily temperature data is not available, so to compute the heat summation index, we will assume an average daily temperature equal to the respective average monthly temperature.

As this is a global dataset, the first step was to clip one of the layers to the extent of California. Then, we compute the heat summation index by using the Raster Calculator to compute the degree days for each month and sum over April through October. We can limit this to the extent of the clipped layer as shown below. Note that we are subtracting 10° instead of 50° because this data follows the centigrade scale.



Also notice that we divide by 100 at the end. This is because the Raster to Vector operation is very computationally expensive, and this will reduce the range of values that it will need to create polygons for by a factor of 100. Though this means our final data will be in chunks of 100 (e.g. a value of 25 reflects an index range of 2450-2549 due to floating-point rounding), this granularity is sufficient. Finally, we vectorize this data to prepare it for overlay analysis alongside soil data.

3.2 Soil Data

To obtain soil data for California, the national Soil Survey Geographic Database (SSURGO) was used. Unfortunately, the data included is organized in a very complicated way. After some investigation through the metadata, I was able to find out which of the 68 tables I needed, and clean the raw text file by adding headers and changing the delimitation symbol. The result was a properly formatted CSV file showing soil information for each specified region in California.

The table join took some time, however. The relationship between the table and the shapefile was many-to-one, because each feature (soil survey area) in the shapefile had a mixture of soil types. Fortunately, the table included percent coverage of soil types for each survey area. This means that I could simply classify each survey area with the soil type that proportionally covers the greatest portion of that area. For instance, if in a specific survey area, 40% of the soil was "well drained", 30% was "excessively drained", and 30% was "poorly drained", then it was classified as having "well drained" soil. To perform this join, it was necessary to sort the CSV data first by soil survey area (using the key), and then by percent coverage. When the join was performed using the key as the join attribute, I ensured that ArcMap took the first row that matched – i.e., the soil type that covers the greatest portion of that survey area. The data is now ready for mapping.

4. Map Creation

4.1 Heat Summation Index Map (Map 1)

From the vector layer created as described in section 3.1, we first used the field calculator in ArcMap to multiply the field that we vectorized by 100, since we previously divided by 100 to vectorize it. We then also need to multiply by 1.8, as the heat summation index is more readily available in Fahrenheit and our data was obtained in centigrade (1°C = 1.8°F).

To make this a little more manageable, a dissolve operation was done to reduce the number of features from 95,547 to 71. By using the standard heat summation index table from section 2, a map showing each of the 5 heat regions throughout California was produced (see map #1).

4.2 Soil Drainage Levels Map (Map 2)

Creating the soil drainage map was straightforward since the relevant data was already joined in section 3.2. The outcome is as expected, with the southeast desertic region of the state showing more drainage compared to the more temperate Central Valley and San Francisco Bay areas (see map #2)

4.3 Combining Temperature and Soil Drainage (Map 3)

Since the purpose of this project is to identify the "best" places for vineyards, we need to somehow combine the maps and data from sections 4.1 and 4.2 into a single viewable product. The logical way to do this, given that the data display entirely different units, is to create a "suitability index" for each factor based on my research. As previously described in section 2, the most desirable temperature conditions are those in Region II, with a heat summation index between 2500 and 3000. Also, the ideal soil drainage condition is the 5th level, "well-drained." As we get farther from these criteria, the spatial suitability for grapevines would decrease. As such, we can create the following scales:

Axel Malahieude
UID 204915219
7/31/2020

Heat Summation Index	Suitability Index	Soil Drainage Level	Suitability Index
< 1000	1	Excessively drained	3
1000 - 1500	2	Somewhat excessively drained	4
1500 - 2000	3	**Well drained**	**5**
2000 - 2500	4	Moderately well drained	4
2500 - 3000	**5**	Somewhat poorly drained	3
3000 - 3500	4	Poorly drained	2
3500 - 4000	3	Very poorly drained	1
4000 - 4500	2	Null values (bodies of water)	0
> 4500	1		

We rate the ideal conditions with a suitability index of 5, and as we get farther, we reduce the suitability index until we get values of 1 or 0 for conditions where grapevines would struggle the most to grow. Now, we can intersect the layers with these two indices and sum to get values ranging between 1 and 10 for a holistic look at how suitable various California regions are for grapevines, where a score of 10 is ideal for winegrowing.

First, we need to attach a suitability index value out of 5 to the two datasets. This is done by exporting the attribute table from each shapefile as a CSV, computing the index in a new column using an Excel formula, and then joining this new data back with the original shapefile. The field calculator in ArcMap was not used because more complicated conditional logic was involved to compute the suitability index.

The Excel formula used for computing the heat summation suitability index:

```
=IF(B2>4500, 1, IF(B2>4000, 2, IF(B2>3500, 3, IF(B2>3000, 4, IF(B2>2500, 5,
IF(B2>2000, 4, IF(B2>1500, 3, IF(B2>1000, 2, 1))))))))
```

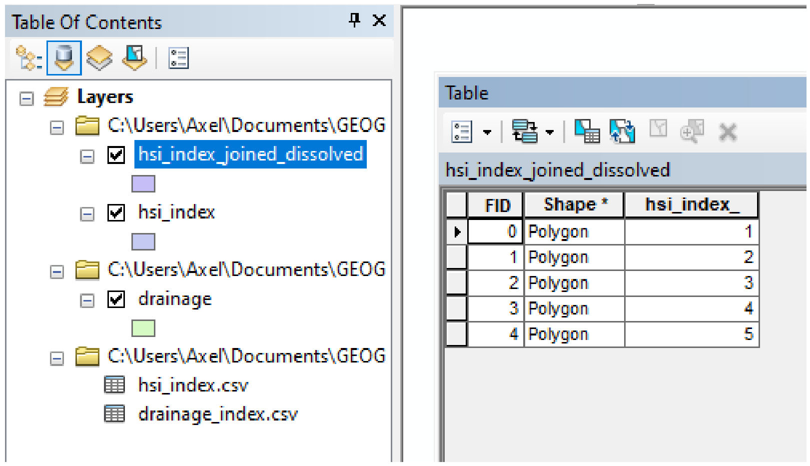
The Excel formula used for computing the soil drainage level suitability index:

```
=IF(D2="excessively drained", 3, IF(D2="somewhat excessively drained", 4,
IF(D2="well drained", 5, IF(D2="moderately well drained", 4, IF(D2="somewhat
poorly drained", 3, IF(D2="poorly drained", 2, IF(D2="very poorly drained",
1, 0)))))))
```

After rejoining these tables to their original source using the FID attribute provided by ArcMap as the key, it is then necessary to dissolve both shapefiles with respect to this new suitability index. The result is a heat summation suitability index shapefile with 5 features, and a soil drainage level suitability index shapefile with 6 features.

Here is what the attribute table looks like after dissolving the heat summation index data, to illustrate:

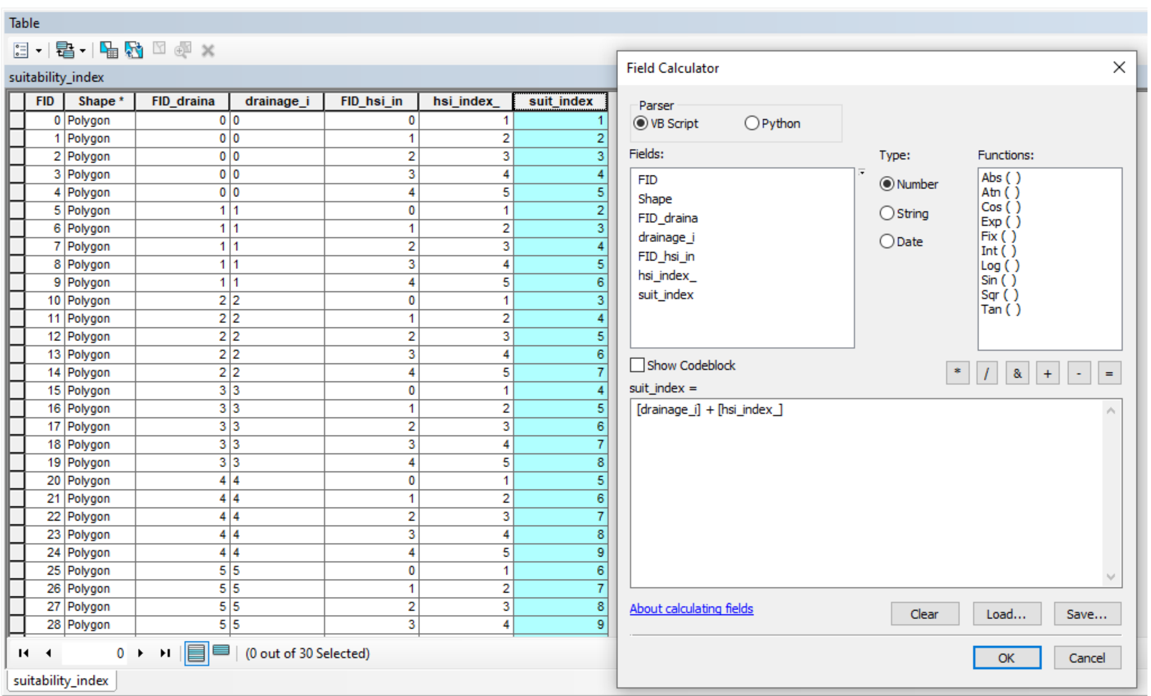
Axel Malahieude
UID 204915219
7/31/2020



FID	Shape *	hsi_index_
0	Polygon	1
1	Polygon	2
2	Polygon	3
3	Polygon	4
4	Polygon	5

Finally, we can intersect these two dissolved layers, yielding a new shapefile with 30 features.

Using the field calculator on this newly created layer, we compute a new overall suitability index with the sum of each existing suitability index value. The result is a value from 1 to 10 associated with each of the 30 features, representing how suitable each area is for grapevines according to these two criteria.



FID	Shape *	FID_draina	drainage_i	FID_hsi_in	hsi_index_	suit_index
0	Polygon	0	0	0	1	1
1	Polygon	0	0	1	2	2
2	Polygon	0	0	2	3	3
3	Polygon	0	0	3	4	4
4	Polygon	0	0	4	5	5
5	Polygon	1	1	0	1	2
6	Polygon	1	1	1	2	3
7	Polygon	1	1	2	3	4
8	Polygon	1	1	3	4	5
9	Polygon	1	1	4	5	6
10	Polygon	2	2	0	1	3
11	Polygon	2	2	1	2	4
12	Polygon	2	2	2	3	5
13	Polygon	2	2	3	4	6
14	Polygon	2	2	4	5	7
15	Polygon	3	3	0	1	4
16	Polygon	3	3	1	2	5
17	Polygon	3	3	2	3	6
18	Polygon	3	3	3	4	7
19	Polygon	3	3	4	5	8
20	Polygon	4	4	0	1	5
21	Polygon	4	4	1	2	6
22	Polygon	4	4	2	3	7
23	Polygon	4	4	3	4	8
24	Polygon	4	4	4	5	9
25	Polygon	5	5	0	1	6
26	Polygon	5	5	1	2	7
27	Polygon	5	5	2	3	8
28	Polygon	5	5	3	4	9

This screenshot shows the results of the intersect operation, and how we can sum two fields to get a final suitability index value.

This suitability index was then colored and mapped to create map #3.

Axel Malahieude
UID 204915219
7/31/2020

5. Comparing Results to Current Winery Locations in California (Maps 4-6)

At this point, we now have a breakdown of how suitable various parts of California are for growing grapes for winemaking. The final step in this project is to compare these results to current winery locations and see how well they did. Firstly, I found a point shapefile on ArcHub of 609 wineries in California. This is incomplete but can give us a general idea as to the spatial extent of California wineries.

Displaying these points atop the suitability map created in the previous section worked, but it became hard to tell what type of location wineries were in because of how large the scale had to be to fit California. To zoom in, data-driven pages were used. Using a shapefile of the 3 main regions (Northern, Central, and Southern California) from the California State Geoportal, 3 maps were created, allowing greater detail in each. In order to display only one region in each map, intersect operations had to be used between the regions and both winery locations and suitability zones.

To further simplify this map, suitability indices were grouped in pairs such that only 5, instead of 10, distinct subdivisions are displayed. For example, zones that scored 9 and 10 were grouped together, as were those that scored 7 and 8, and so on. Maps 4-6 show the results, displaying Northern, Central, and Southern California respectively with the wineries located in those regions.

6. Conclusions and Reflections

In conclusion, most wineries actually chose their locations very well, residing mostly in regions that are rated 9 or 10 on this suitability index. That means that they have soil that drains relatively well and favorable seasonal temperatures. However, comparing Southern California to the other two regions, we see that wineries are typically in zones that scored 7 or 8, which is a whole tier below wineries in the rest of California.

Suitability Index	Number of Wineries
10	79
9	247
8	113
7	147
6	17
5	2
4	1
3	0
2	0
1	0

Based on these maps, the California coast north of Los Angeles up to near the Oregon border is the ideal location for vineyards. The foothills of the Sierra Nevada also score very well, but the Central Valley and southeast desert should be avoided.

Axel Malahieude
UID 204915219
7/31/2020

While conclusions are generally consistent with what one would expect (after all, wineries would likely do research into good locations for their vineyards), there are a few anomalies and limitations to note. Firstly, the two variables analyzed in this project are not the only ones to consider. In addition, though a region might be classified as having seasonal temperatures too high for most grape varieties (e.g. Southern California), placing a vineyard on a gently sloping hill, in a valley with enough wind, or at a high enough altitude may result in favorable conditions (Goldammer). In other words, though a potential vineyard site may score low on this suitability index, it may actually be good based on its microclimate.

With regards to this project, it was harder than expected to find factors affecting grapevines relevant to California because in general, conditions are very favorable. Considerations that are made on the East Coast or in Europe such as mean days of sunlight, frost potential, and aspect (which way slopes face in order to get the most of the sunlight) are not often considered in California since California typically benefits from good sun exposure and infrequent snows or frosts.

For future projects, it would be interesting to extend this project to look at other famous winemaking regions such as France or Italy, and how they compare to California. In addition, perhaps more factors such as sunlight days and rainfall could be included into the suitability index that was created in this project in order to get an even better idea of how regions compare in their vineyard potential.

Axel Malahieude
UID 204915219
7/31/2020

References

Amerine, M. A. and A. J. Winkler. "Composition and Quality of Musts and Wines of California Grapes." *Hilgardia* (1944): 184. Document.

CalWineries. *UC Davis Heat Summation Scale*. 21 April 2020. 27 July 2020. <https://www.calwineries.com/learn/grape-growing/climate/heat-summation-scale>.

Goldammer, Ted. *Grape Grower's Handbook: A Guide to Viticulture for Wine Production*. Centreville: APEX Publishers, 2018.

Horiuchi, Gladys. "California Wine Sales Reach $43.6 Billion in U.S. Market in 2019." 9 July 2020. *Wine Institute.* 27 July 2020. <https://wineinstitute.org/press-releases/california-wine-sales-reach-43-6-billion-in-u-s-market-in-2019/>.

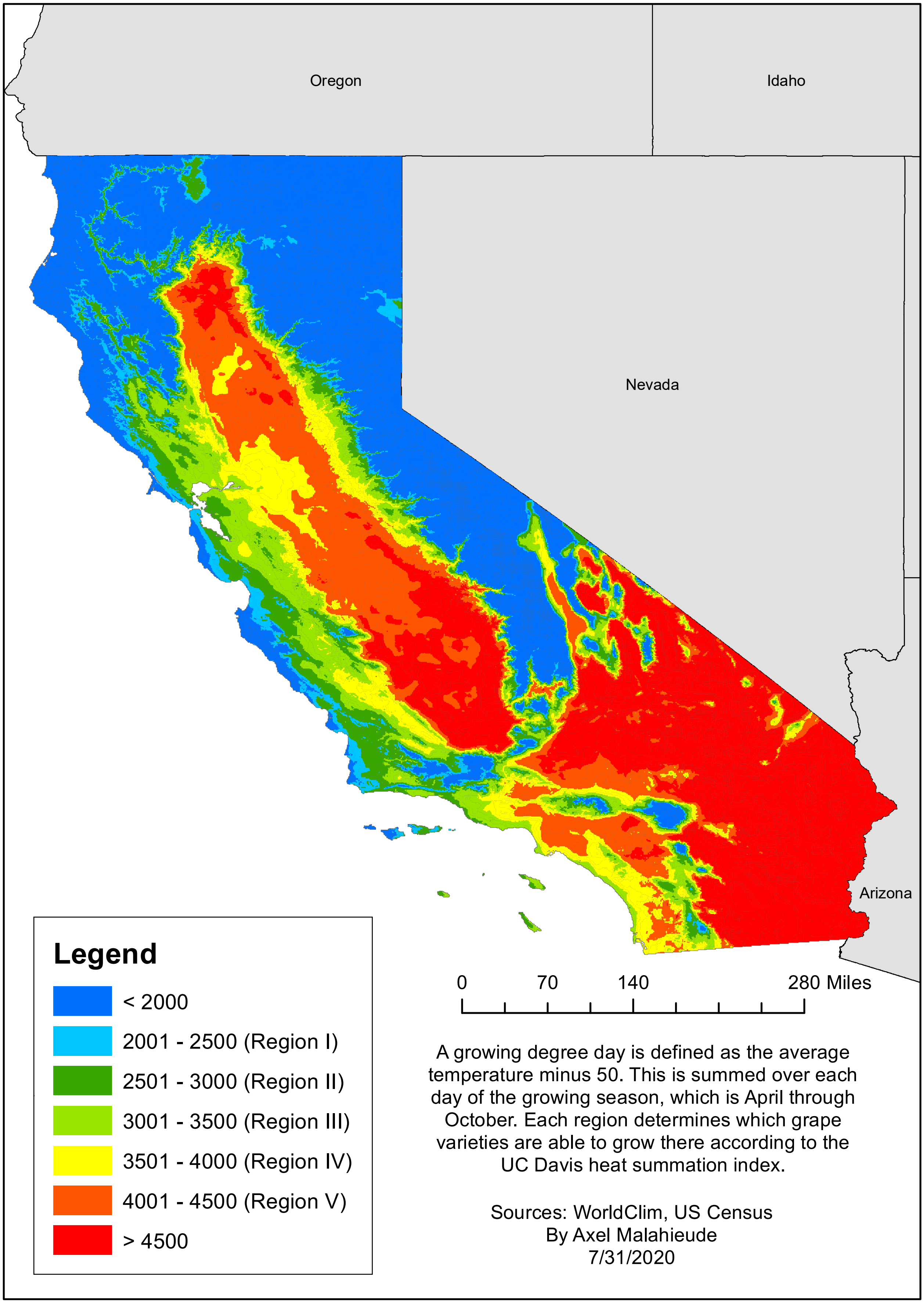
Johnson, Hugh and Jancis Robinson. *The World Atlas of Wine*. London: Octopus Publishing Group, 2013. Print.

Soil Survey Division Staff. *Drainage Class*. Soil Survey Manual. Washington, D.C.: U.S. Department of Agriculture, 1993. Document.

Data Sources:

Soil quality, SSURGO Web Soil Survey: https://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm
Historical temperature data, World Climate: https://www.worldclim.org/data/worldclim21.html
State boundaries, US Census: https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html
State regional boundaries, CA Geoportal: https://gis.data.ca.gov/datasets/DWR::dwr-region-offices
California wineries: https://www.arcgis.com/home/item.html?id=61b983d24b58455590d5ec88f8d9efaa

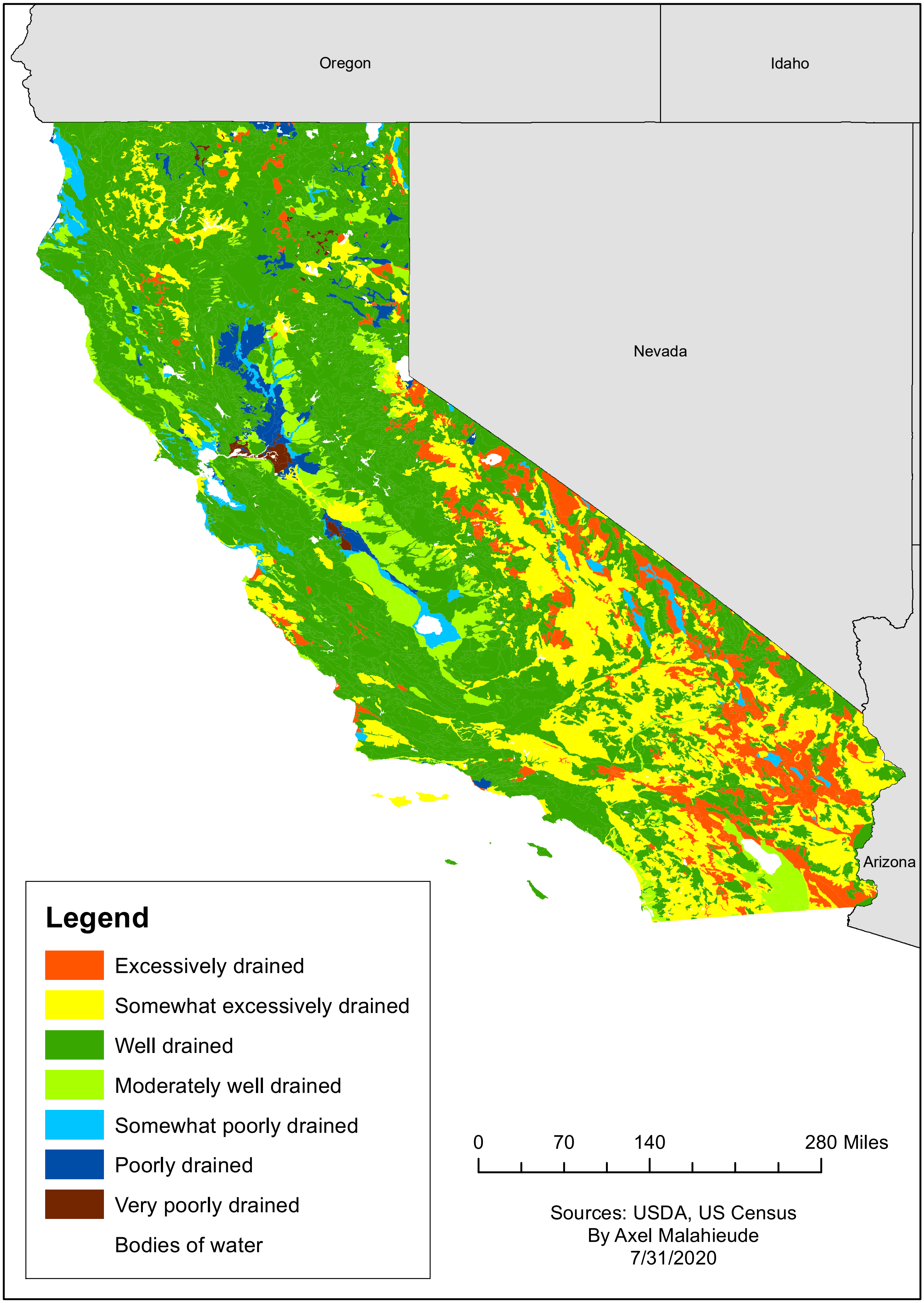
Grapevine Growing Degree Days in California



A growing degree day is defined as the average temperature minus 50. This is summed over each day of the growing season, which is April through October. Each region determines which grape varieties are able to grow there according to the UC Davis heat summation index.

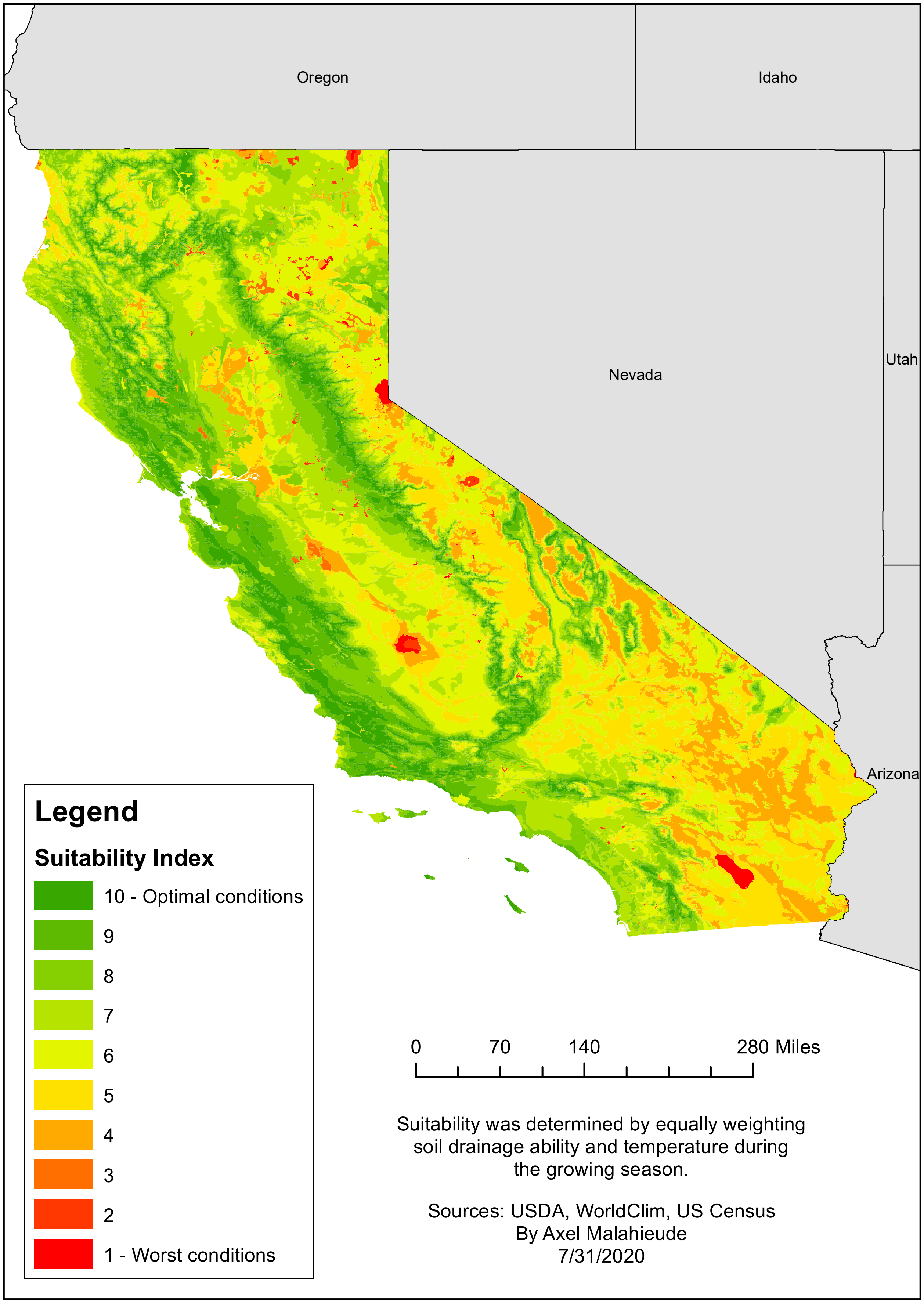
Sources: WorldClim, US Census
By Axel Malahieude
7/31/2020

USDA Soil Drainage Levels Across California



Sources: USDA, US Census
By Axel Malahieude
7/31/2020

Overall Grapevine Suitability of California

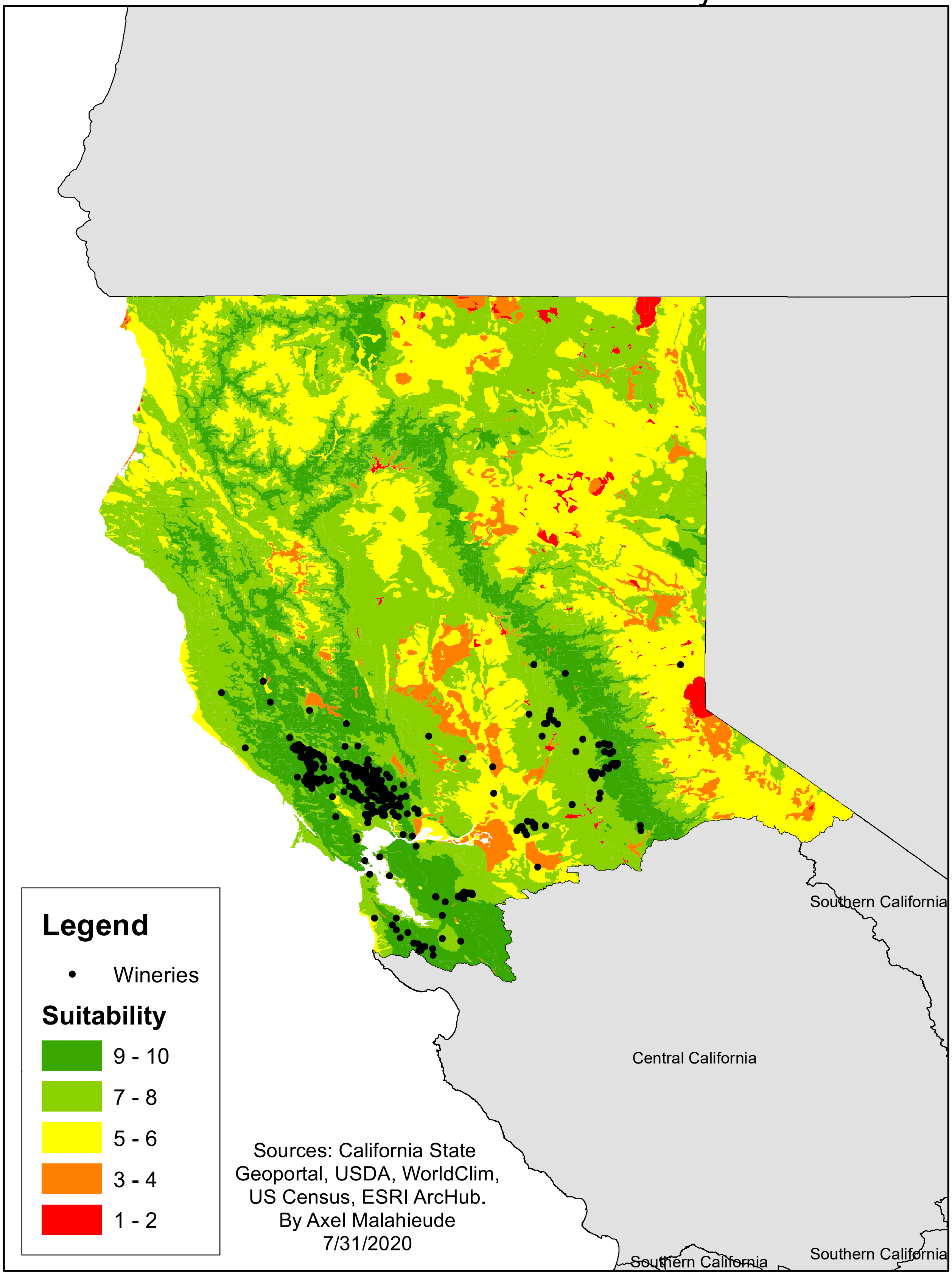


Suitability was determined by equally weighting soil drainage ability and temperature during the growing season.

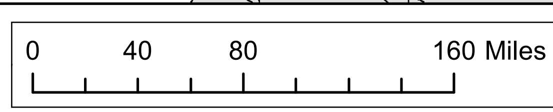
Sources: USDA, WorldClim, US Census
By Axel Malahieude
7/31/2020

Northern California

Winery Site Selection

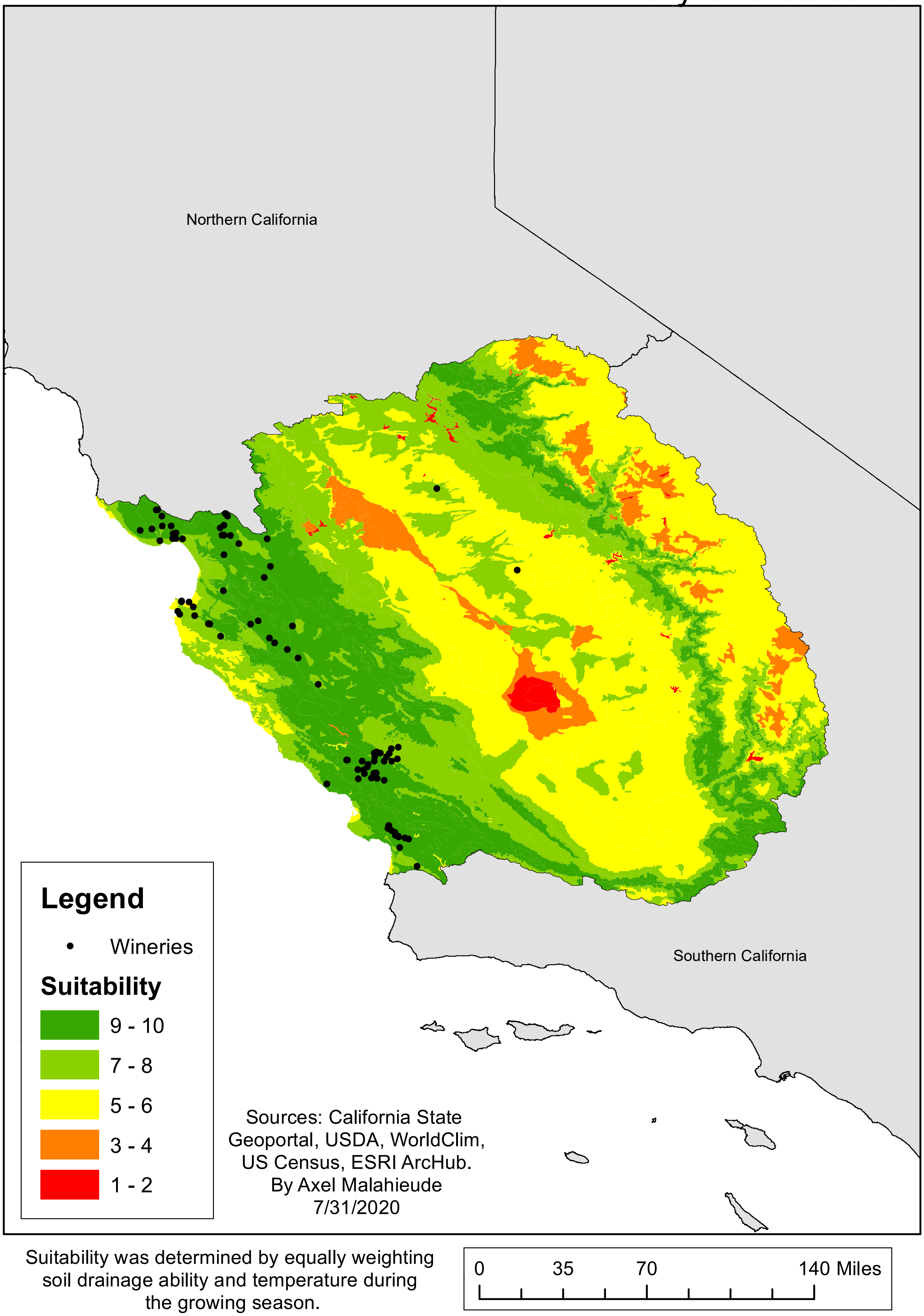


Suitability was determined by equally weighting soil drainage ability and temperature during the growing season.



Central California

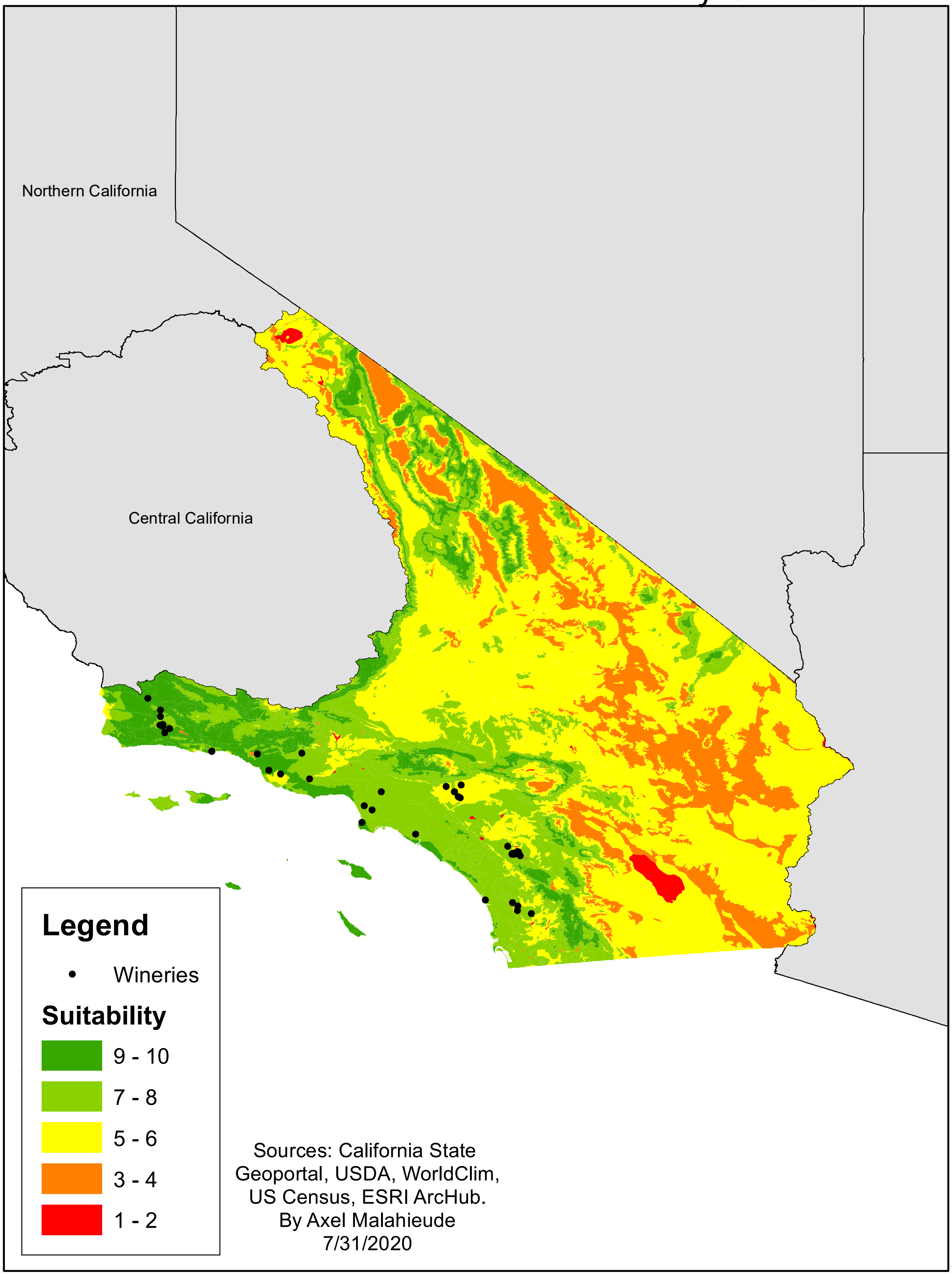
Winery Site Selection



Suitability was determined by equally weighting soil drainage ability and temperature during the growing season.

Southern California

Winery Site Selection



Suitability was determined by equally weighting soil drainage ability and temperature during the growing season.

0 50 100 200 Miles