Instructions for Calculating Societal Environmental Variables from their Domesticates’ Niche Trait Data

See the accompanying “domesticate\_niche\_data.xlsx” spreadsheet, which walks through the following analysis:

* **Step 1**: open “combined.csv”
* **Step 2**: Separate out the 12 rows pertaining to animals from the 37 pertaining to plants (column AH: H = animals, A = plants)
* **Step 3**: Calculate the “COUNT” presence for each society-column of all animal rows: e.g. SCCS10 = COUNT presence: 2 (this is the number of different animal species present under each society)
* **Step 4**: Count the animal species for every society that show the code “4” or “1” and add these to the total “COUNT” of animal species for each society, we will call this the weighted COUNT:
  1. e.g. “=SUM(COUNTIF(H2:H13,4),COUNTIF(H2:H13,1),H16)”
  2. Essentially every animal presence of “4” or “1” has been counted twice. (This is because these are primary species, while animals with the “2” code are secondary species – primary species will be weighted twice as heavily – thus the denominator for calculating MEAN values for each society will include “4”s and “1”s twice.)
* **Step 5**: For every society create a new range of cells (below) for all the animal species. Into this new array the Environmental Variable value for each animal species will be copied, if that animal shows a “4”, “1” or “2”, indicating its presence in that society. However, the Environmental variable value will be doubled if the presence code is for a primary domesticate (code “4” or “1”).
  1. e.g. “=SUM(IF(H2=4,$DQ2\*2,),IF(H2=1,$DQ2\*2,),IF(H2=2,$DQ2))”
  2. The “DQ” in the above example refers to the column of the Environmental Variables, these will all be replaced by the columns of the other Environmental Variables (in **Step 7**) to give values average values for each society, for all 41 Environmental Variables.
  3. The above formula can be copied into the cells for every domesticated animal species, for every society.
* **Step 6**: Take the (weighted (“4” and “1” = \*2, “2” = \*1)) SUM for the Environmental Variable values of all domesticated animals for each society, and divide it by the weighted COUNT (from **Step 4)** of each society.
  1. “=(SUM(H19:H30))/H17”
* **Step 7**: **Steps 3 to 6** give the Partially-Weighted animal domesticate values for the first Environmental Variable.
  1. Copy these values for each society into a new table, under “Animal, *Environmental Variable Name*”. (Note: Pasting with the “Values Only” option will prevent the code behind these numbers being copied across and altering the output values)
  2. Then change the column code in the formula for **Step 5**, to move onto the next Environmental Variable. (Note: By selecting all cells with the relevant formula and using Find/ Replace, you can change all column values simultaneously, e.g. “DU” to “DV”)
  3. Repeat this step 41 times for all Environmental Variables, copying the results each time into the new table.
* **Step 8**: Now, within the Plant data (separated in **Step 2**), separate out all the Buckwheat, Wheat, Millet, Yam and Banana sub-species, as these have all been derived from a single occurrence report within D-PLACE, and must therefore be averaged within themselves before a value can be committed to the Plant average for each society.
* **Step 9**: For Buckwheat:
  1. First count the number of different sub-species of Buckwheat that are being averaged for each society (e.g. “=COUNT(I37:I38)”, where I37:I38 is the presence for each species).
  2. Then follow a similar principle as shown in **Step 5**, allocating a new cell for each of the Buckwheat sub-species. However, as all of the different sub-species are weighted equally within each society, there is no need, at this point, to add weightings; this is simply an average of the Buckwheat sub-species’ values for each environmental variable. (“=SUM(IF(I37=4,$DS37),IF(I37=1,$DS37),IF(I37=2,$DS37))” – in every cell of the grid for each Buckwheat sub-species, for every society).
  3. Finally, calculate the average of the Environmental variable values for those sub-species of Buckwheat that were present in each society (e.g. code: “=IF(I40<1,0,(SUM(I42:I43)/I40))”). (Note: the use of “IF(I40<1,0,” is to avoid errors arising from dividing by 0 on societies with 0 Buckwheat sub-species present.)
* **Step 10**: Repeat **Step 9** for Wheat, Millet, Yam and Banana.
* **Step 11**: Take the remaining 18 plant species rows, and append to the end of this data an absence/ societal weighting row for each of the five averaged plant species (Buckwheat, Wheat, Millet, Yam and Banana).
  1. To do this, use the code “=IF((MAX(I37:I38))<1,"NA",(MAX(I37:I38)))”, for each averages species, and for every society, where the cells “I37:I38” are the weightings for every sub-species within each of the five averaged species. (As all sub-species occurrences within the same society were weighted equally, we can just take the “MAX” value or “NA”.)
* **Step 12**: Now take the 23 plant species (the unaltered 18, and the 5 averaged species: **Steps 9-11**):
  1. Calculate a plant *Count Presence* row (see **Step 3**).
  2. Calculate a plant *Weighted Counts* row (see **Step 4**).
* **Step 13**: Calculate a plant species *Weighted Environmental Variable* array, for the 18 unaltered plant species (as in **Step 5**).
  1. Then add to this array *Weighted Environmental Variables* for the five averaged plant species. To do this the code must be altered somewhat, as these values are not on the same rows as each species, instead un-weighted averages were calculated above in **Step 9iii**. Instead use “=SUM(IF(I128=4,I45\*2,),IF(I128=1,I45\*2,),IF(I128=2,I45))” where the averaged Environmental variable for each species is take from the **Step 9iii** result for each society.
* **Step 14**: Then calculate the average of all domesticated plant species’ weighted environmental variables for each society (see **Step 6**).
* **Step 15**: Repeat **Step 7** for the plant species data, pasting into the same table as with the animal data, under “Plant, *Environmental Variable Name*”, thus generating a table of 82 rows (two for every environmental variable), with a column for each society.
* **Step 16**: Locate the “relative\_weightings\_H-A.csv” (animal Husbandry: Agriculture) table and multiply every cell belonging to the animal data by the “H” (animal Husbandry) weighting value for each society and every value related to the plant environmental data, by the “A” (Agriculture) weighting value for each society.
  1. It is useful to transpose the table from **Steps 7&15** here, so that the left-most column shows the society, with the environmental variables and whether they are plant (A) or animal (H) derived, in the top rows. The formula “=H258\*$C258” or “=I258\*$D258” can then be used, with the “$C...” and “$D...” alternating for the animal then plant Environmental Variable contribution value for each society.
  2. Note: Some societies (SCCS: 153, 163, 28, 34, 46, 61 and 65) have no domesticates in one category (plants/ animals), for these societies, we take exclusively the value given for the type of domesticates the society do present, without any input from the listed weighting in the “Weightings H:A” table (their weighting is obviously just “1” for their respective domesticate-category). (Note, for these societies the formula can simply copy the values from the table built is Steps 7&15, as no alteration need take place).
* **Step 17**: Sum the animal and plant values for each society, for every environmental variable. (To avoid errors resulting from attempting to sum cells that contain the “#DIV/0!” message, use the formula: “=SUMIF(H374:I374,">-999999999999999999999")”.) This is the final weighted value for that society (see the final sheet of the accompanying “combined.xlsx”).

The results of this calculation can be found on the final sheet of the “domesticate\_niche\_data.xlsx” document. They are also saved in .csv format in the file “societal\_niche\_traits.csv”.