OPTIONS FORMCHAR="|----|+|---+=|-/\<>\*";

\* if processes get too slow, run this to free up memory, then rerun relevant

sections;

\* proc datasets library=work kill noprint; **run**;

\*import herb data;

**proc** **import** datafile="C:\Users\Emily\Desktop\herbx1.csv"

out=herbx dbms=csv replace; getnames=yes; **run**; \* N = 122547;

\* creating a set of herbs;

**data** herb1x; set herbx;

\* removing blank lines;

if sspp=' ' then delete;

\* removing 1999--data are of extremely poor quality;

if year = **1999** then delete;

\* type 1--it is a plant. type 2--zero plants were found in that plot/year;

\*type = 1;

\*if (sspp = 'XXXXx') then type = 2;

keep aspect bcat coun quad covm elev hydrn plot slope soileb sspp year;

rename coun=count soileb=soil hydrn=hydr covm=cover;

**run**; \* n = 12,543;

**proc** **sort** data=herb1x; by sspp plot quad year bcat cover count soil elev slope aspect hydr; **run**;

\* proc print data=herb1x (firstobs=1 obs=20); title 'herb1x'; **run**;

/\*

\* getting stem counts to look at;

proc sort data=herb1x; by sspp; run;

proc means data=herb1x noprint n sum mean min max; by sspp; var count;

output out=sumstems n=n sum=sumcount mean=meancount min=mincount max=maxcount;

data sumstems1; set sumstems; drop \_TYPE\_ \_FREQ\_; RUN;

proc sort data=sumstems1; by sumcount n;

\* proc print data=sumstems1; title 'sumstems';run;

\*\*\*\*\*\*\*\*\*includes sspp, n, sumcount, meancount, mincount, maxcount;

\*5 species w/ highest stem count: DILI2, DIOLx, HELA5, DISP2, POPR4;

\*/

\* work only top 5 species;

**data** fivesp; set herb1x; if (sspp='DILI2' | sspp='DIOLx' | sspp='HELA5' | sspp='DISP2' | sspp='LETEx');

\* proc print data=fivesp (firstobs=1 obs=10); title 'fivesp'; **run**; \* n = 2908;

\*\*\*\*\*\*\*\*\*includes all vars for 5 species;

\* --- plot translation dataset--orig plot names to nums 1-56;

**data** plotid; set fivesp; dummy = **1**; keep plot dummy;

**proc** **sort** data=plotid; by plot; **run**;

**proc** **means** data=plotid noprint mean; by plot; var dummy;

output out=plotid2 mean = mean;

\* proc print data=plotid2; title 'plotid2'; **run**;

**data** plotid3; set plotid2; plotnum = \_n\_; keep plot plotnum;

\* proc print data=plotid3; title 'plotid3';

**run**; \* n = 54, max = 55;

\*\*\*\*\*\*\*\*\*includes plot and plotnum;

\*55 herbx plots, 54 fivesp plots.

missing plot 1224--herbs were counted once in 2006, none of these species appeared;

\* --- species translation dataset--orig sp codes to nums 1-315;

**proc** **sort** data=fivesp; by sspp;

**proc** **means** data=fivesp mean noprint; var count; by sspp;

output out=splist mean=mcoun;

**data** splist2; set splist; spnum=\_n\_; keep sspp spnum;

\* proc print data=splist2; title 'splist2'; **run**;

\*\*\*\*\*\*\*\*\*includes sspp and spnum;

\*----- merge into one data set -----------------;

\* merge in species codes;

**proc** **sort** data=fivesp; by sspp;

**proc** **sort** data=splist2; by sspp;

**data** step1; merge fivesp splist2; by sspp;

\* merge in plots;

**proc** **sort** data=step1; by plot;

**proc** **sort** data=plotid3; by plot;

**data** fivesp2; merge step1 plotid3; by plot;

if (year < **2012**) then yearnum = **1**;

if (year > **2011**) then yearnum = year - **2010**;

\*\*\* fivesp2 has all 17 variables, including eviro vars, plotnum, spnum, yearnum \*\*\*\*\*\*\*\*;

**run**; \* N = 2908;

/\*

proc print data=fivesp2 (firstobs=1 obs=30); title 'fivesp2'; \*var quad plot plotnum;

run;

\*/

\* ---- get plot data for environmental variables, to use later;

**proc** **sort** data=fivesp2; by plotnum yearnum;

**proc** **means** data=fivesp2 mean noprint; by plotnum plot yearnum;

var bcat cover soil elev slope aspect hydr;

output out=plotvars mean=bcat cover soil elev slope aspect hydr;

**data** plotvars2; set plotvars;

keep plotnum yearnum bcat cover soil elev slope aspect hydr;

**run**; \* N = 207;

\* ----- fix up counts for multiple obs in a single quad-sp-year;

\* the fix up for pre and post fire is different;

**proc** **sort** data=fivesp2; by plotnum quad spnum yearnum;

**data** prefire; set fivesp2; if yearnum = **1**; \* n = 213;

**data** postfire; set fivesp2; if yearnum > **1**; \* n = 2695;

**run**;

\*proc print data=postfire (firstobs=1 obs=30); title 'postfire'; **run**;

\* average multiple counts of same quad-sp-year in prefire years;

**proc** **means** data=prefire mean noprint; by plotnum quad spnum yearnum;

var count; output out=prefiremeans mean = count;

\*proc print data=prefiremeans; title 'prefiremeans'; **run**; \* n = 149;

\* recombine to make one fixed data set of counts;

**data** fivesp3; set prefiremeans postfire;

keep plotnum quad spnum yearnum count; **run**; \* n = 2844;

\*----- create numerical data set for iml, called fivesp3 -------------;

**proc** **sort** data=fivesp3; by plotnum quad spnum yearnum;

\*proc print data=fivesp3 (firstobs=1 obs=30); title 'fivesp3'; **run**; \*n=2844;

\*proc contents data=fivesp3; **run**;

\* order in fivesp3 is plotnum, quad, spnum, yearnum, count;

/\*

\*finding plots that were not visited each year (to be used later, post-iml, to remove 0 counts that

should be missing values;

proc sort data=fivesp3; by plotnum yearnum;

proc means data=fivesp3 mean noprint; by plotnum yearnum;

var count;

output out=fivesp3x mean=mcount;

run;

proc print data=fivesp3x; run;

proc freq data=fivesp3x; table plotnum\*yearnum; run;

\*/

\*------------------IML------------------------------------------;

**proc** **iml**;

use fivesp3; read all into matin; \* print matin;

nrowsmatin = nrow(matin); \* print nrowsmatin;

nplots = **54**;

nquadsperplot = **10**;

nspecies = **5**;

nyrs = **5**; \* assumes we will treat pre-fire as a single year;

nrowsmatout = nquadsperplot \* nplots \* nspecies \* nyrs;

matcountquad = j(nrowsmatout,**6**,**0**);

maxrowsperplot = nquadsperplot\*nspecies\*nyrs;

maxrowsperquad = nspecies\*nyrs; maxrowspersp = nyrs;

\*print maxrowsperplot, maxrowsperquad, maxrowspersp;

\* set up matcountquad with one row per sp x quad x year;

\* cols: col1:row# col2: plot col3:uniquad col4:sp col5:year col6:count;

holdplot=**1**; holdquad=**1**; holdsp=**1**; holdyr=**1**;

do i = **1** to nrowsmatout;

matcountquad[i,**1**] = i;

matcountquad[i,**2**] = holdplot;

matcountquad[i,**3**] = holdquad;

matcountquad[i,**4**] = holdsp;

matcountquad[i,**5**] = holdyr;

holdyr = holdyr + **1**;

if holdyr = **6** then do; holdyr = **1**; holdsp = holdsp + **1**; end;

if holdsp = **6** then do; holdsp = **1**; holdquad = holdquad + **1**; end;

if holdquad = **11** then do; holdquad = **1**; holdplot = holdplot + **1**; end;

end;

\* print (matcountquad[1:30,]);

\*input: 1-plot 2-quad 3-sp 4-year 5-count;

\*output: 1-rowid 2-plot 3-quad 4-sp 5-year 6-count;

do i = **1** to nrowsmatin; \* going line by line through input matrix;

\* get info from input matrix;

tempplot = matin[i,**1**];

tempquad = matin[i,**2**];

tempsp = matin[i,**3**];

tempyr = matin[i,**4**];

tempcount = matin[i,**5**];

\* calculate the output matrix target rows;

outrow = (tempplot-**1**)\*maxrowsperplot + (tempquad-**1**)\*maxrowsperquad + (tempsp-**1**)\*maxrowspersp + (tempyr);

\* if tempplot=1 then do;

\* print tempsp,tempyr,tempcount;

\* end;

\* write to outrow;

matcountquad[outrow,**1**] = outrow;

matcountquad[outrow,**2**] = tempplot;

matcountquad[outrow,**3**] = tempquad;

matcountquad[outrow,**4**] = tempsp;

matcountquad[outrow,**5**] = tempyr;

matcountquad[outrow,**6**] = tempcount;

end;

\* print (matcountquad[1:30,]);

\* print matcountquad;

cols = {rowid plotnum quad spnum yearnum count};

create imlout1 from matcountquad [colname = cols];

append from matcountquad;

**quit**;

**run**;

\* ----------output data row = plot-quad-sp-year ----------------;

\*proc print data=imlout1 (firstobs=1 obs=30); title 'imlout1'; **run**; \*n=13500;

\*proc print data=imlout1 (firstobs=13400 obs=13500); title 'imlout1'; **run**; \*n=13500;

\*proc contents data=imlout1; **run**;

\*rowid, plotnum, quad, spnum, yearnum, count;

\*----- merge back into one data set -----------------;

\* merge back in species codes;

**proc** **sort** data=imlout1; by spnum;

**proc** **sort** data=splist2; by spnum;

**data** temp1; merge imlout1 splist2; by spnum; **run**;

\*proc sort data=temp1; \*by rowid;

\*proc print data=temp1 (firstobs=1 obs=30); title 'temp1'; **run**;

\* merge back in plots;

**proc** **sort** data=temp1; by rowid plotnum quad spnum yearnum;

**proc** **sort** data=plotid3; by plotnum;

**data** temp2; merge temp1 plotid3; by plotnum; **run**;

\*proc print data=temp2 (firstobs=1 obs=30); title 'temp2'; **run**;

\* merge back in environmental vars;

**proc** **sort** data=temp2; by plotnum yearnum;

**proc** **sort** data=plotvars2; by plotnum yearnum;

**data** herbbyquad; merge temp2 plotvars2; by plotnum yearnum;

if ((yearnum=**1**) and (plotnum=**19** | plotnum=**40** | plotnum=**41** | plotnum=**42** | plotnum=**43** | plotnum=**44** |

plotnum=**45** | plotnum=**46** | plotnum=**47** | plotnum=**48** | plotnum=**49** | plotnum=**50** | plotnum=**51** |

plotnum=**52** | plotnum=**53** | plotnum=**54**)) then count=**.** and pa=**.**;

if ((yearnum=**2**) and (plotnum=**4** | plotnum=**5** | plotnum=**6** | plotnum=**7** | plotnum=**8** | plotnum=**12** |

plotnum=**14** | plotnum=**15** | plotnum=**20** | plotnum=**21** | plotnum=**24** | plotnum=**25** | plotnum=**26** |

plotnum=**27** | plotnum=**29** | plotnum=**30** | plotnum=**31** | plotnum=**33** | plotnum=**34** | plotnum=**35** |

plotnum=**38** | plotnum=**39** | plotnum=**46**)) then count=**.**;

if ((yearnum=**3**) and (plotnum=**24** | plotnum=**25** | plotnum=**26** | plotnum=**27** | plotnum=**33** | plotnum=**34** |

plotnum=**38** | plotnum=**39**)) then count=**.**;

if ((yearnum=**4**) and (plotnum=**24** | plotnum=**25** | plotnum=**26** | plotnum=**27** | plotnum=**33** | plotnum=**34** |

plotnum=**38** | plotnum=**39**)) then count=**.**;

if ((yearnum=**5**) and (plotnum=**24** | plotnum=**25** | plotnum=**26** | plotnum=**27** | plotnum=**33** | plotnum=**34** |

plotnum=**38** | plotnum=**39**)) then count=**.**;

if count=**0** then pa=**0**; if count>**0** then pa=**1**;

**run**;

\*proc print data=herbbyquad (firstobs=1050 obs=1080); title 'herbbyquad'; **run**;

/\*

proc contents data=herbbyquad; run;

\*rowid, plotnum, wuad, spnum, yearnum, count, all others;

proc print data=herbbyquad;

var plot plotnum quad sspp spnum yearnum count bcat soil pa;

title 'herbbyquad'; run;

\*/

\* USE HERBBYQUAD FOR ANALYSES BY QUAD;

\* ------ use iml to re-arrange: all yrs in a row-------------------------------;

**data** foriml2; set herbbyquad; keep plotnum quad spnum yearnum count cover;

\*proc print data=foriml2; title 'foriml2'; **run**;

\*proc contents data=foriml2; **run**;

\*1--plotnum, 2--quad, 3--spnum, 4--yearnum, 5--count, 6--cover;

**proc** **iml**;

use foriml2; read all into matcountquad1; \* print matcountquad1;

nrowsmatcountquad1 = nrow(matcountquad1); \*print nrowsmatcountquad1;

nplots = **54**;

nquadsperplot = **10**;

nspecies = **5**;

nyrs = **5**;

\*print nplots, nquadsperplot, nspecies, nyrs;

\* creating a new matrix with a column for each year,

and a row for each plot-quad-species;

\* 14 columns: rowid plotnum quad spnum year1 year2 year3 year4 year5

cover1 cover2 cover3 cover4 cover5;

ncolsmatcountquad2 = **14**;

nrowsmatcountquad2 = nplots \* nquadsperplot \* nspecies;

matcountquad2=j(nrowsmatcountquad2,ncolsmatcountquad2,**0**);

newmaxrowsperplot = nquadsperplot\*nspecies; \* 10x5 = 50; \*print newmaxrowsperplot;

newmaxrowsperquad = nspecies; \* should be 5; \*print newmaxrowsperquad;

holdplot=**1**; holdquad=**1**; holdsp=**1**;

do i = **1** to nrowsmatcountquad2;

matcountquad2[i,**1**] = i;

matcountquad2[i,**2**] = holdplot;

matcountquad2[i,**3**] = holdquad;

matcountquad2[i,**4**] = holdsp;

holdsp = holdsp + **1**;

if holdsp = **6** then do; holdsp = **1**; holdquad = holdquad + **1**; end;

if holdquad = **11** then do; holdquad = **1**; holdplot = holdplot + **1**; end;

end;

\* print matcountquad2;

\*1--plotnum, 2--quad, 3--spnum, 4--yearnum, 5--count, 6--cover;

do i = **1** to nrowsmatcountquad1; \* going line by line through previously created matrix;

\* get info from input matrix;

tempplot = matcountquad1[i,**1**];

tempquad = matcountquad1[i,**2**];

tempsp = matcountquad1[i,**3**];

tempyr = matcountquad1[i,**4**];

tempcount = matcountquad1[i,**5**];

tempcov = matcountquad1[i,**6**];

\* calculate the new matrix col;

newcolcount = tempyr + **4**;

newcolcover = tempyr + **9**;

\* calculate the new matrix target rows;

newrow = (tempplot-**1**)\*newmaxrowsperplot + (tempquad-**1**)\*newmaxrowsperquad + tempsp;

matcountquad2[newrow,**1**] = newrow;

matcountquad2[newrow,**2**] = tempplot;

matcountquad2[newrow,**3**] = tempquad;

matcountquad2[newrow,**4**] = tempsp;

matcountquad2[newrow,newcolcount] = tempcount;

matcountquad2[newrow,newcolcover] = tempcov;

\* if tempplot=2 then do; \* print tempplot tempquad tempsp tempyr newrow; \* end;

end;

\*print matcountquad2;

\* labeling columns;

countnames2 = {rowid plotnum quad spnum count1 count2 count3 count4 count5

cover1 cover2 cover3 cover4 cover5};

create imlout2 from matcountquad2 [colname = countnames2];

append from matcountquad2;

**quit**;

**run**;

\* proc print data=imlout2 (firstobs=1 obs=100); title 'imlout2'; **run**; \*n=2700;

\* proc print data=imlout2 (firstobs=2601 obs=2700); title 'imlout2'; **run**;

\* proc contents data=imlout2; **run**; \*rowid, plotnum quad, spnum, counts1-5;

\*----- merge back into one data set -----------------;

\* merge back in species codes;

**proc** **sort** data=imlout2; by spnum;

**proc** **sort** data=splist2; by spnum;

**data** temp3; merge imlout2 splist2; by spnum; **run**;

\* merge back in plots & environmental vars;

**proc** **sort** data=temp3; by plotnum;

**proc** **sort** data=plotid3; by plotnum;

**proc** **sort** data=plotvars2; by plotnum;

**data** quadhistory; merge temp3 plotid3 plotvars2; by plotnum;

drop cover yearnum;

if ((count1=**0**) and (plotnum=**19** | plotnum=**40** | plotnum=**41** | plotnum=**42** | plotnum=**43** | plotnum=**44** |

plotnum=**45** | plotnum=**46** | plotnum=**47** | plotnum=**48** | plotnum=**49** | plotnum=**50** | plotnum=**51** |

plotnum=**52** | plotnum=**53** | plotnum=**54**)) then count1=**.**;

if ((count2=**0**) and (plotnum=**4** | plotnum=**5** | plotnum=**6** | plotnum=**7** | plotnum=**8** | plotnum=**12** |

plotnum=**14** | plotnum=**15** | plotnum=**20** | plotnum=**21** | plotnum=**24** | plotnum=**25** | plotnum=**26** |

plotnum=**27** | plotnum=**29** | plotnum=**30** | plotnum=**31** | plotnum=**33** | plotnum=**34** | plotnum=**35** |

plotnum=**38** | plotnum=**39** | plotnum=**46**)) then count2=**.**;

if ((count3=**0** | count4=**0** | count5=**0**) and (plotnum=**24** | plotnum=**25** | plotnum=**26** | plotnum=**27** |

plotnum=**33** | plotnum=**34** | plotnum=**38** | plotnum=**39**)) then count3=**.**;

if ((count4=**0** | count4=**0** | count5=**0**) and (plotnum=**24** | plotnum=**25** | plotnum=**26** | plotnum=**27** |

plotnum=**33** | plotnum=**34** | plotnum=**38** | plotnum=**39**)) then count4=**.**;

if ((count5=**0** | count4=**0** | count5=**0**) and (plotnum=**24** | plotnum=**25** | plotnum=**26** | plotnum=**27** |

plotnum=**33** | plotnum=**34** | plotnum=**38** | plotnum=**39**)) then count5=**.**;

if count1=**0** then pa1=**0**; if count1>**0** then pa1=**1**;

if count2=**0** then pa2=**0**; if count2>**0** then pa2=**1**;

if count3=**0** then pa3=**0**; if count3>**0** then pa3=**1**;

if count4=**0** then pa4=**0**; if count4>**0** then pa4=**1**;

if count5=**0** then pa5=**0**; if count5>**0** then pa5=**1**;

**run**; \*n=2700;

\*proc print data=quadhistory; title 'quadhistory'; **run**;

/\*

proc sort data=quadhistory; by rowid; run;

proc print data=quadhistory (firstobs=1000 obs=1500); title 'quadhistory';

var plotnum pa5 count5 cover5 spnum bcat soil quad; run;

proc contents data=quadhistory; run;

proc sort data=herbbyquad; by rowid; run;

proc print data=herbbyquad (firstobs=1000 obs=1050); title 'herbbyquad';

var plotnum yearnum spnum count pa cover bcat soil quad; run;

proc contents data=herbbyquad; run;

\*/

/\*

proc export data=quadhistory

outfile='Desktop\work2a.csv'

dbms=csv

replace;

run;

\*/