test isOrigin.Oct.9.2018

Chao Ma

Oct. 9, 2018

Install isOrigin package from Github

```
#devtools::install_github("SPATIAL-Lab/isorig", force=T)
```

Load library

```
library(isOrigin)
```

```
## Loading required package: raster
```

```
## Loading required package: sp
```

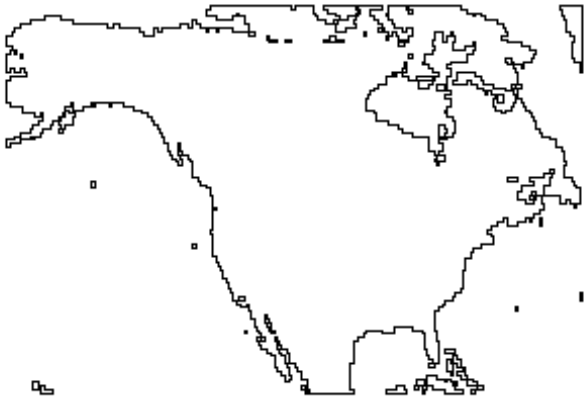
```
## Loading required package: ggplot2
```

```
## Loading required package: rgdal
```

```
## rgdal: version: 1.3-4, (SVN revision 766)
##  Geospatial Data Abstraction Library extensions to R successfully loaded
##  Loaded GDAL runtime: GDAL 2.2.3, released 2017/11/20
##  Path to GDAL shared files: C:/Users/Chao Ma/Documents/R/win-
library/3.5/rgdal/gdal
##  GDAL binary built with GEOS: TRUE
##  Loaded PROJ.4 runtime: Rel. 4.9.3, 15 August 2016, [PJ_VERSION: 493]
##  Path to PROJ.4 shared files: C:/Users/Chao Ma/Documents/R/win-
library/3.5/rgdal/proj
##  Linking to sp version: 1.3-1
```

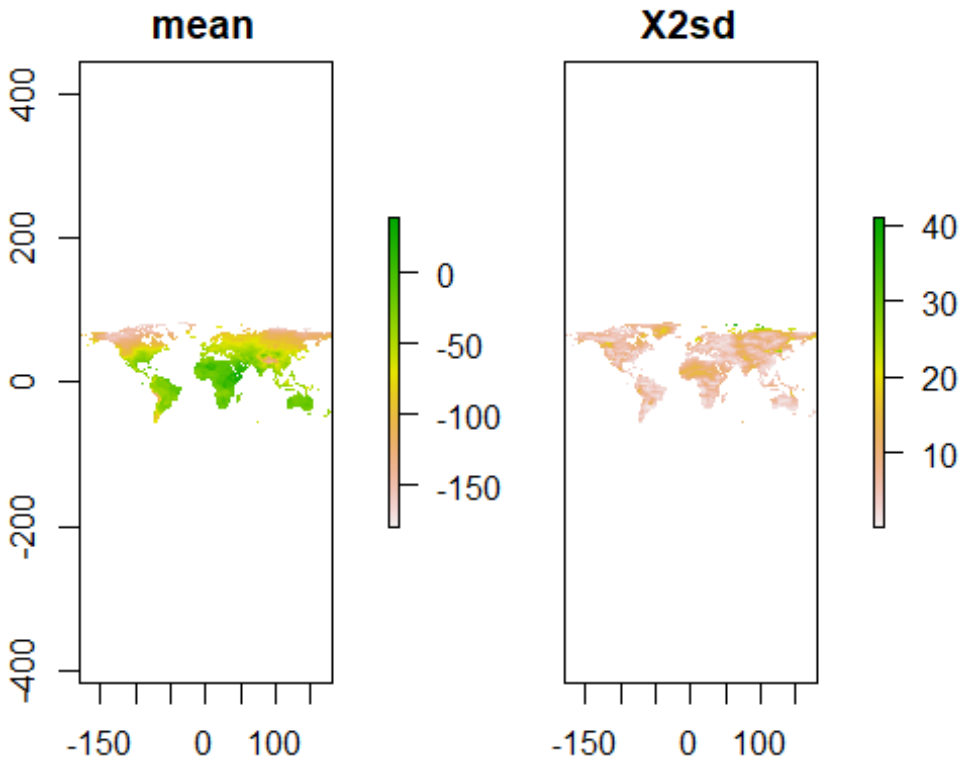
Load North America mask

```
data("naMap")
plot(naMap)
```



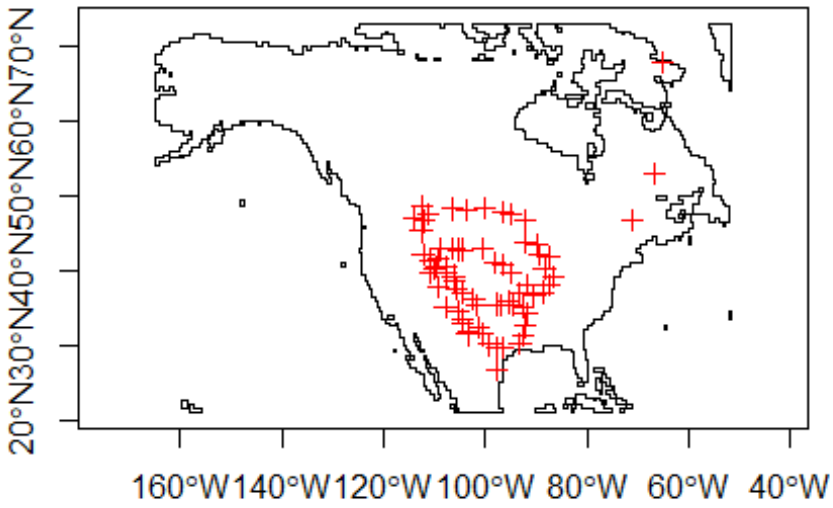
Load world precipitation hydrogen isoscape

```
data("d2h_world")
plot(d2h_world)
```



Load hydrogen isotope for human hair in North America

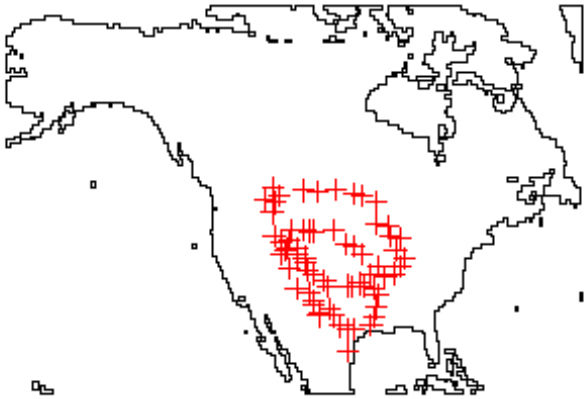
```
d = subOrigData(taxon = c("Homo sapiens"), mask = naMap)
```




```
## 233 data points are found
```

Exclude some outliers. This step is optional, which depends on your data quality

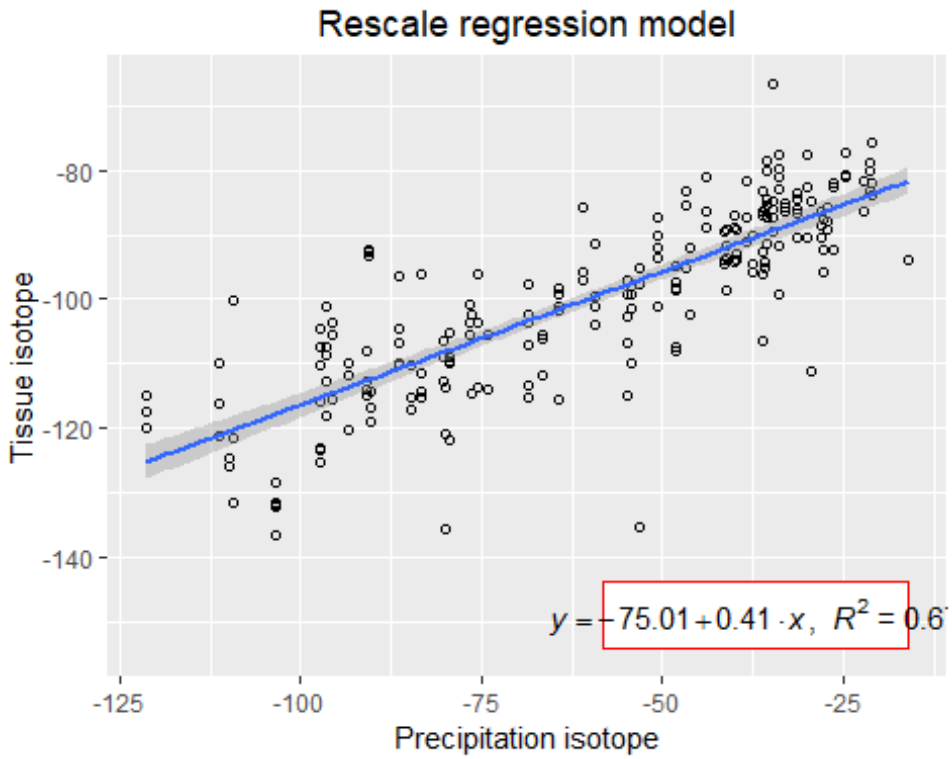
```
d <-as.data.frame(d)
dd = d[d$coords.x1<(-80),]
dd <- SpatialPointsDataFrame(dd[,2:3], as.data.frame(dd[,1]))
crs(dd) <- "+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0"

plot(naMap)
plot(dd, add=T, col=2)
```

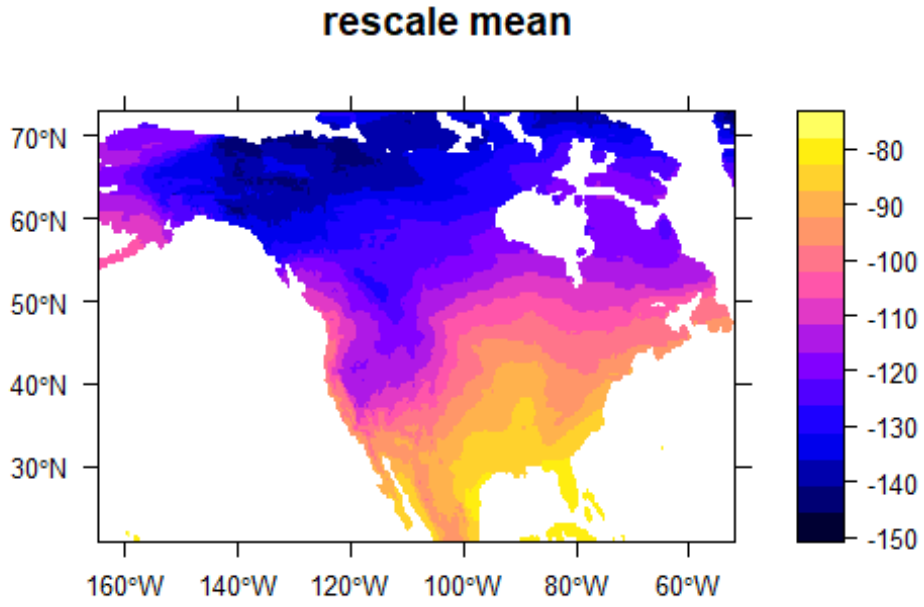


Rescale from environmental isoscape to tissue isoscape

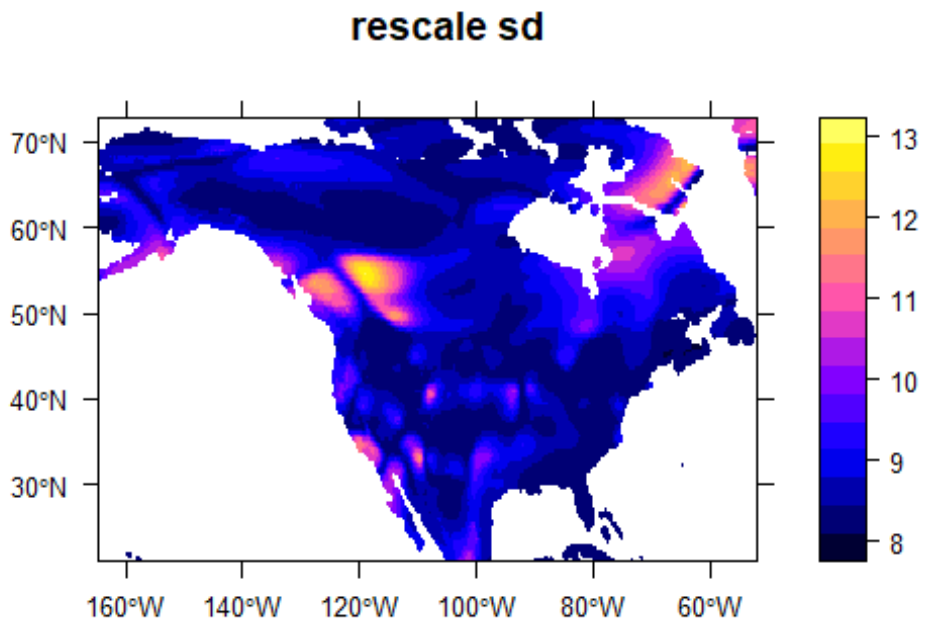
```
r = calRaster(known = dd, isoscape = d2h_world, mask = naMap)

##
##
## ------------------------------------------------------------------------
-------
## rescale function uses linear regression model, the summary of this model
is:
## ------------------------------------------------------------------------
-------
##
## Call:
## lm(formula = tissue.iso ~ isoscape.iso[, 1])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -38.440  -4.029   0.216   4.442  22.779
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -75.00744    1.32392  -56.66   <2e-16 ***
## isoscape.iso[, 1]   0.41440    0.01998   20.74   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 8.078 on 213 degrees of freedom
## Multiple R-squared:  0.6688, Adjusted R-squared:  0.6672
## F-statistic:   430 on 1 and 213 DF,  p-value: < 2.2e-16
```





```
## Warning in dir.create("output"): 'output' already exists
```

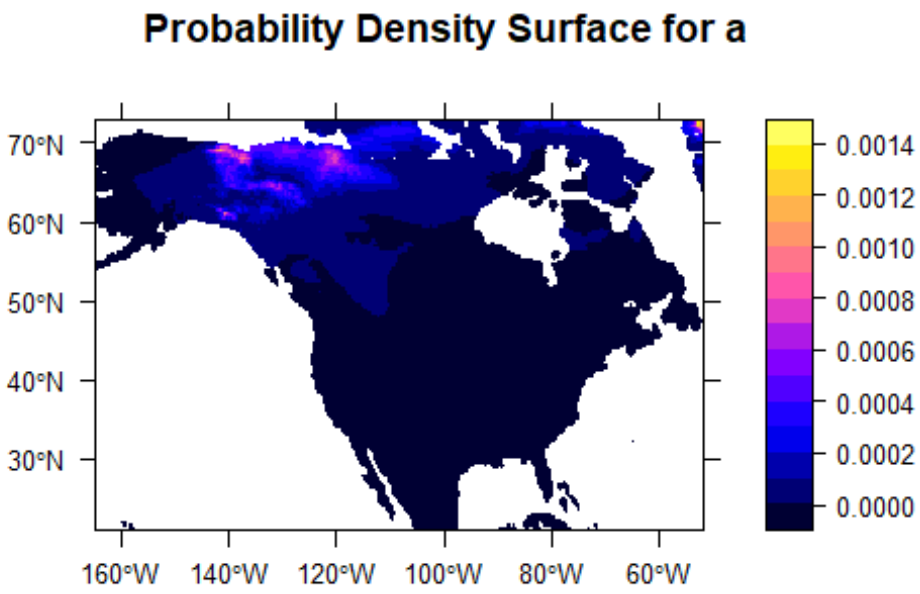


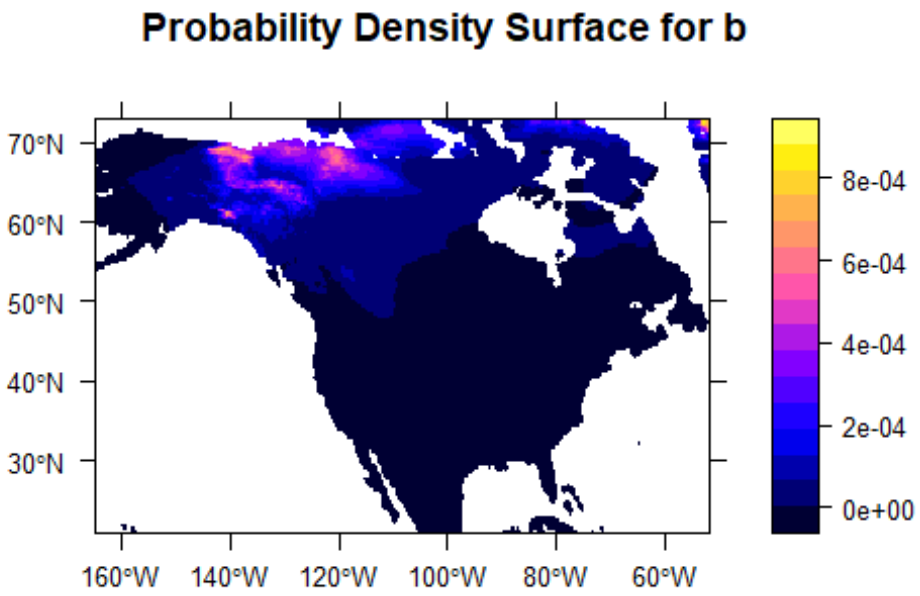
Four unknown-origin examples

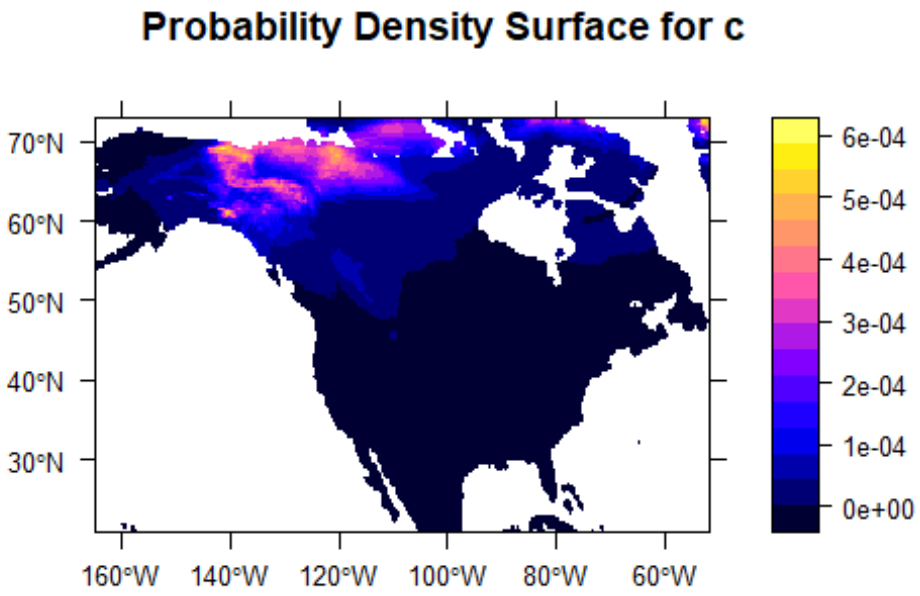
```
id = letters
d2H = seq(-160, -80, by=80/25)
un = data.frame(id,d2H)
```

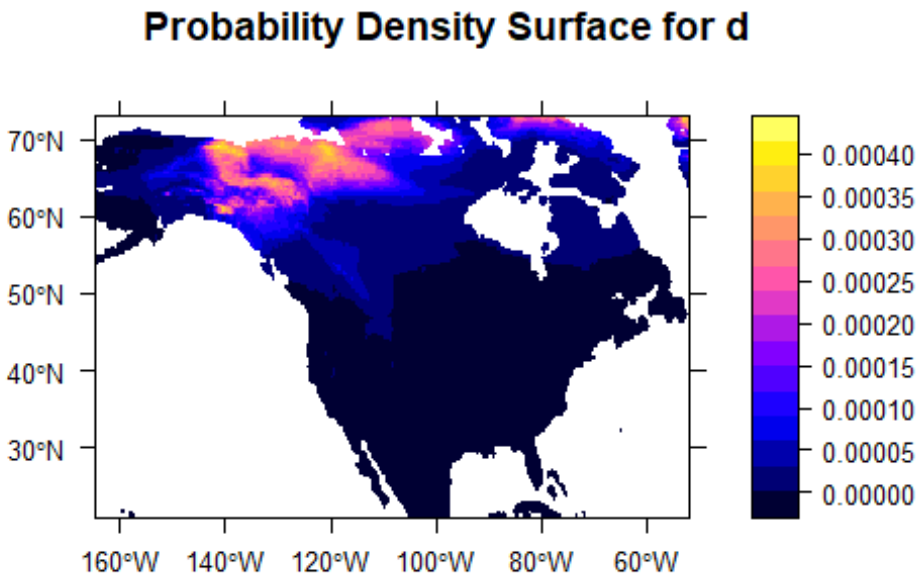
Assignment for unknown-origin examples

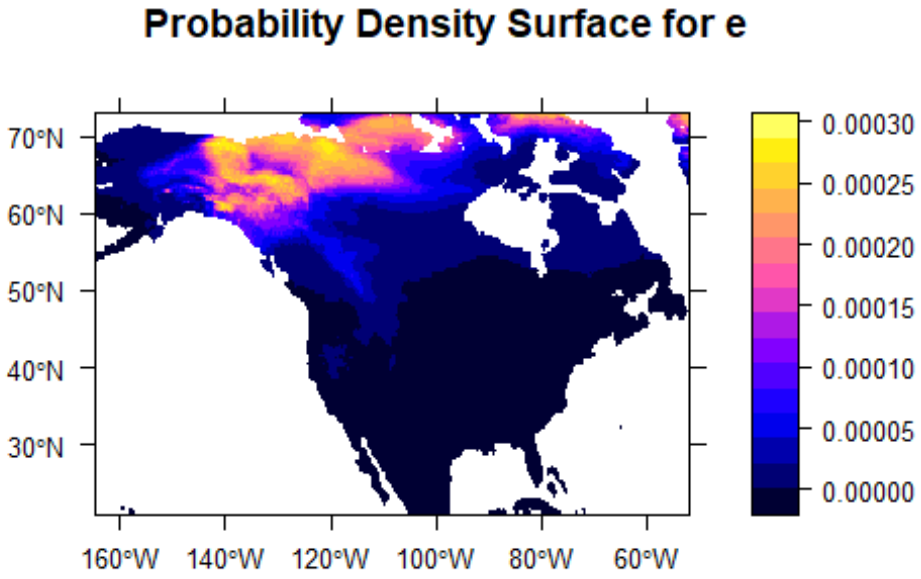
```
asn = pdRaster(r$isoscape.rescale,unknown=un,mask=naMap)

## Warning in dir.create("output"): 'output' already exists

## Warning in dir.create("output/pdRaster_Gtif"): 'output\pdRaster_Gtif'
## already exists
```

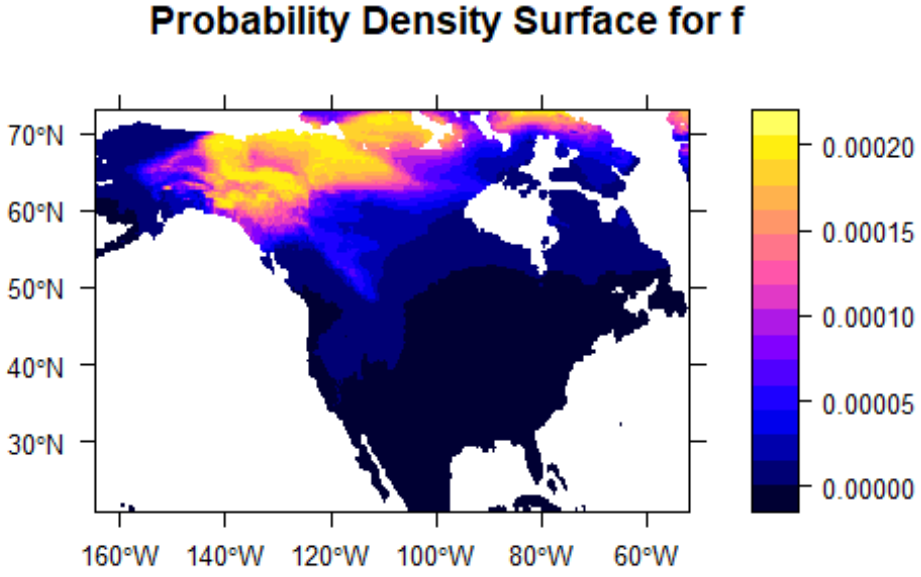

Probability Density Surface for a
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.0014
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000

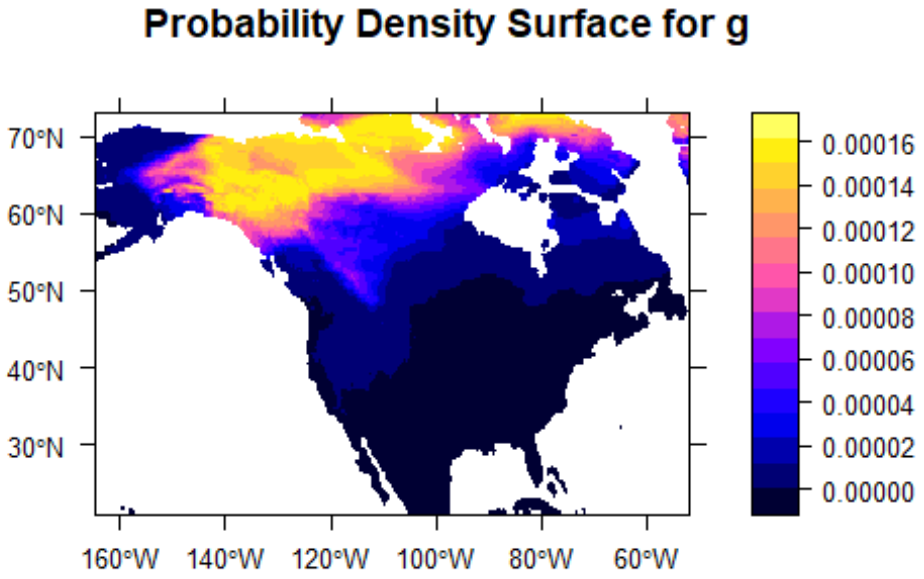

Probability Density Surface for b
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
8e-04
6e-04
4e-04
2e-04
0e+00

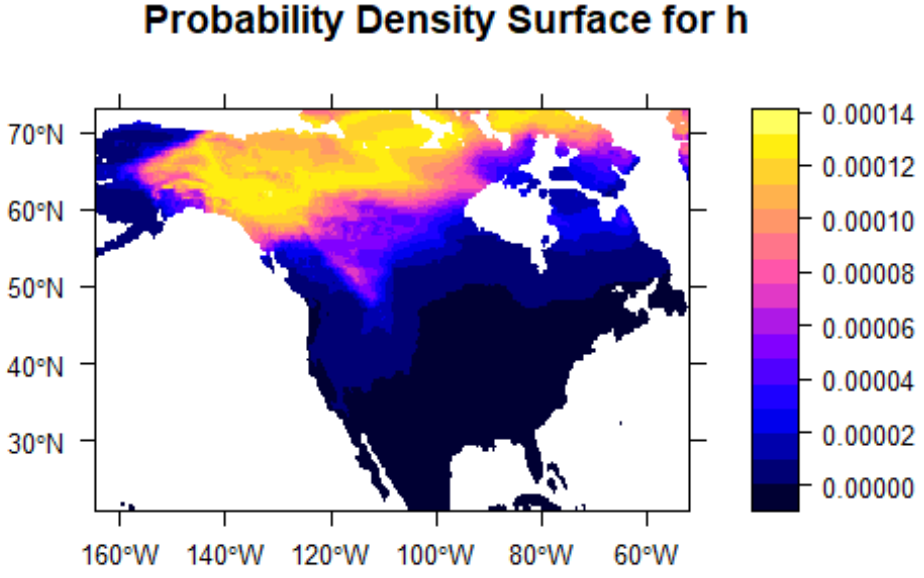

Probability Density Surface for c
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
6e-04
5e-04
4e-04
3e-04
2e-04
1e-04
0e+00

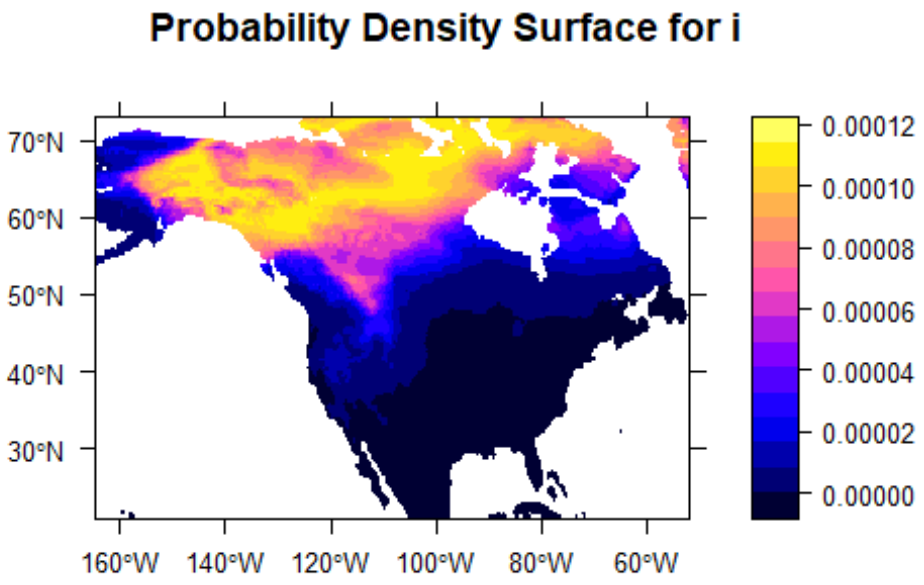

Probability Density Surface for d
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00040
0.00035
0.00030
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000

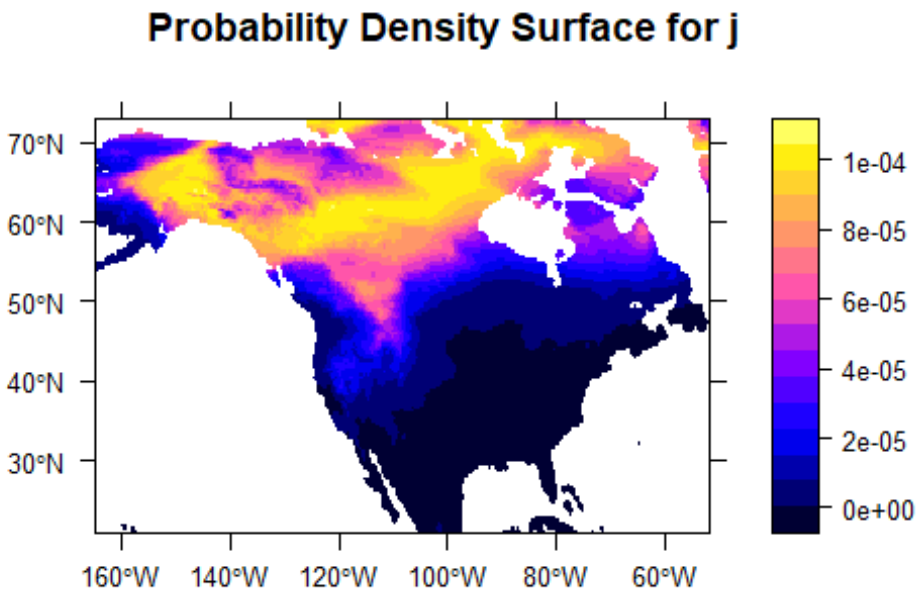

Probability Density Surface for e
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00030
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000

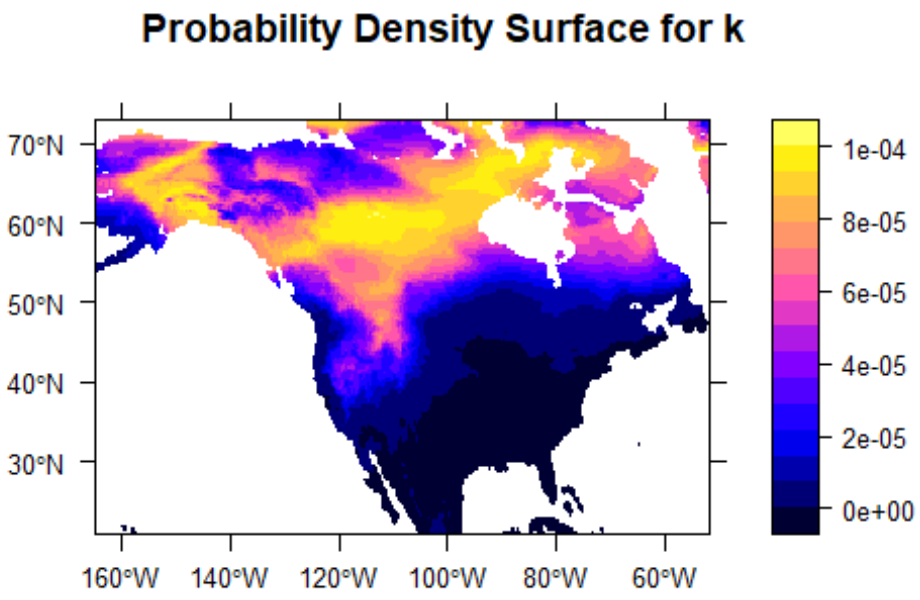

Probability Density Surface for f
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00020
0.00015
0.00010
0.00005
0.00000

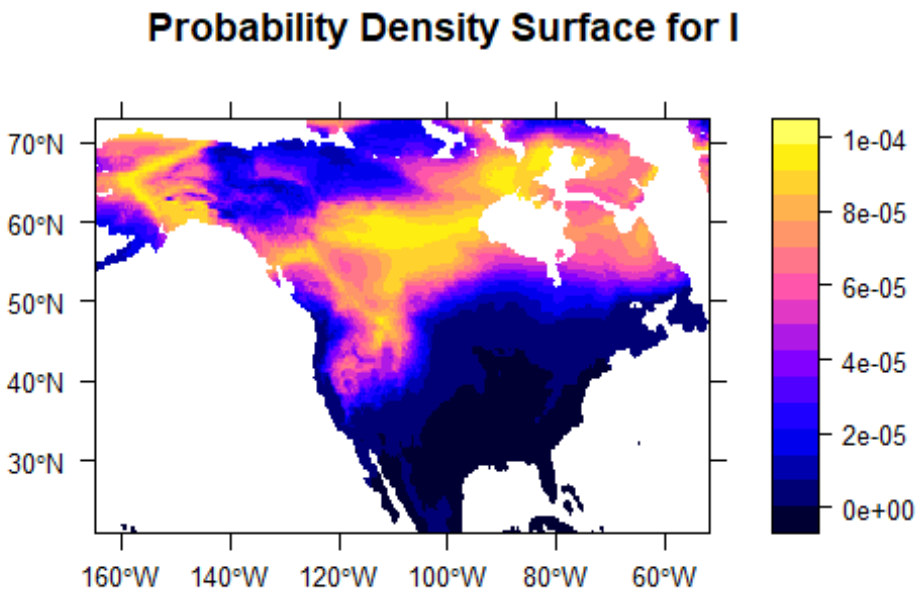

Probability Density Surface for g
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00016
0.00014
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000

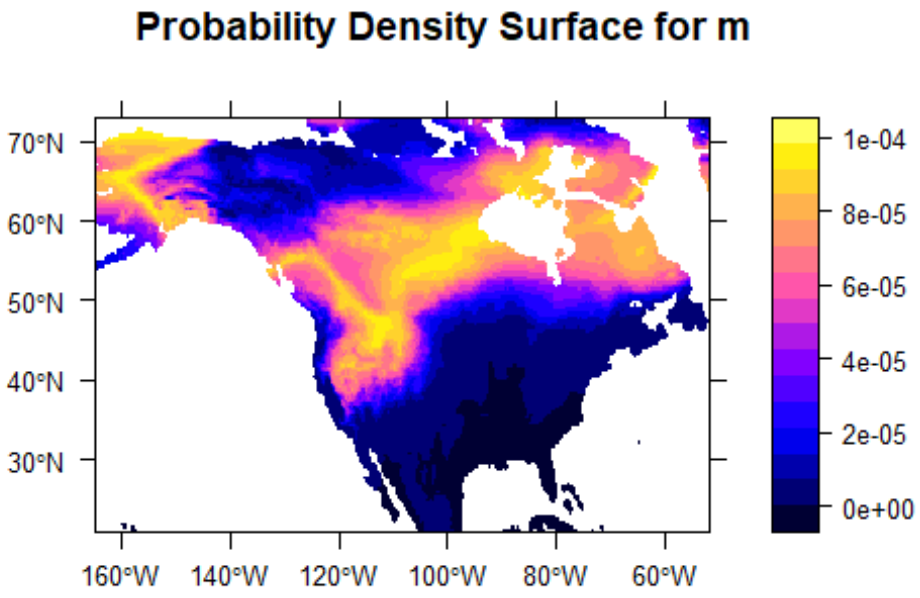

Probability Density Surface for h
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00014
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000

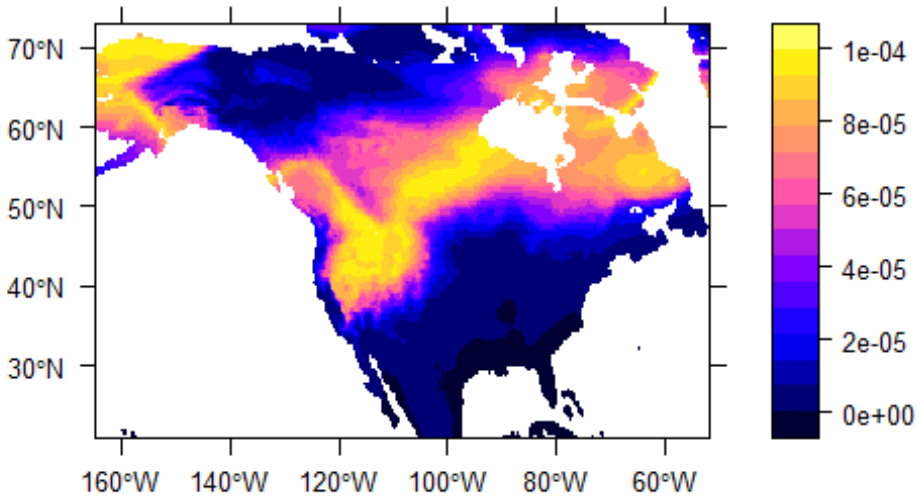

Probability Density Surface for i
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000

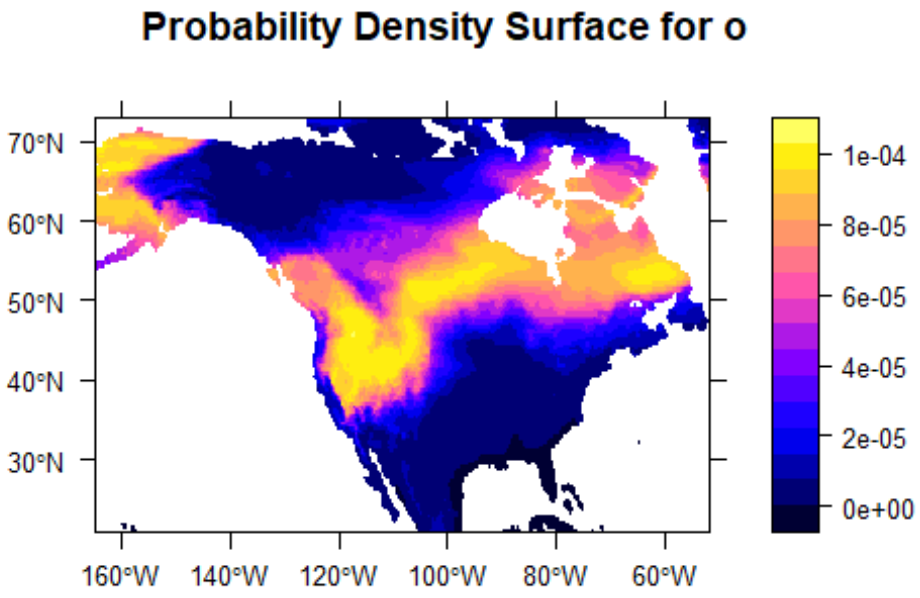

Probability Density Surface for j
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
1e-04
8e-05
6e-05
4e-05
2e-05
0e+00

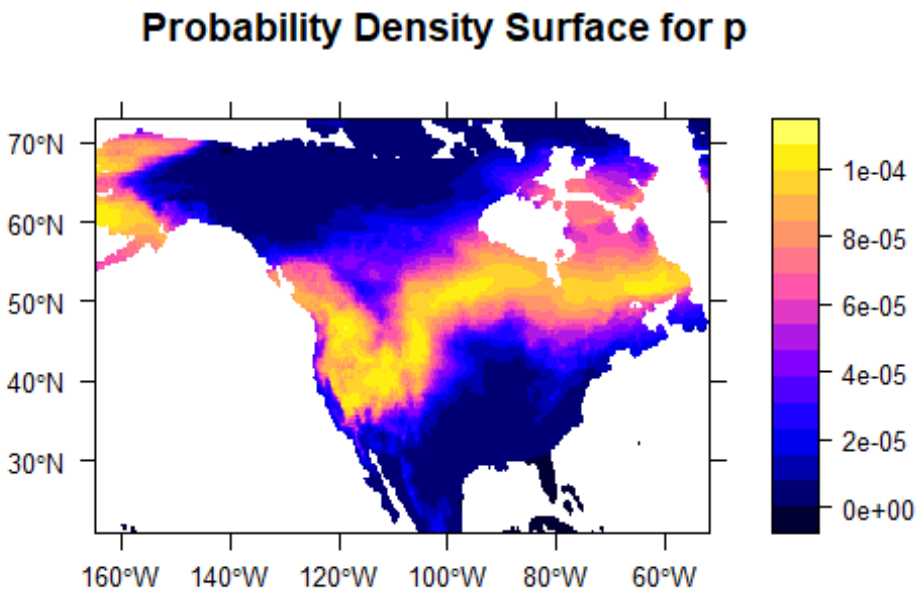

Probability Density Surface for k
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
1e-04
8e-05
6e-05
4e-05
2e-05
0e+00


Probability Density Surface for l
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
1e-04
8e-05
6e-05
4e-05
2e-05
0e+00

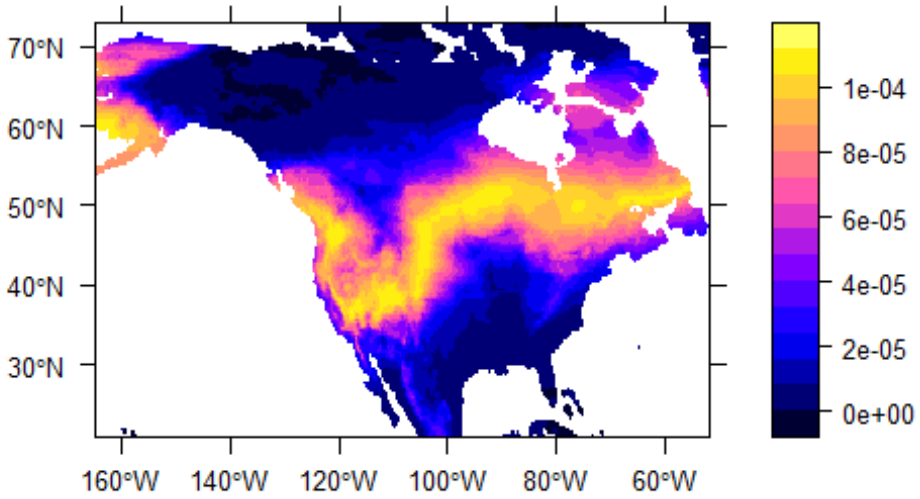

Probability Density Surface for m
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
1e-04
8e-05
6e-05
4e-05
2e-05
0e+00


Probability Density Surface for n
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
1e-04
8e-05
6e-05
4e-05
2e-05
0e+00

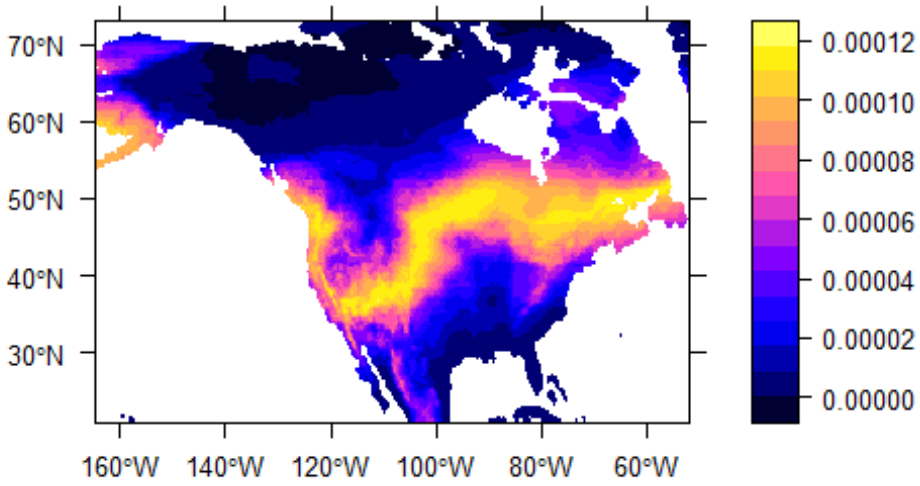

Probability Density Surface for o
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
1e-04
8e-05
6e-05
4e-05
2e-05
0e+00

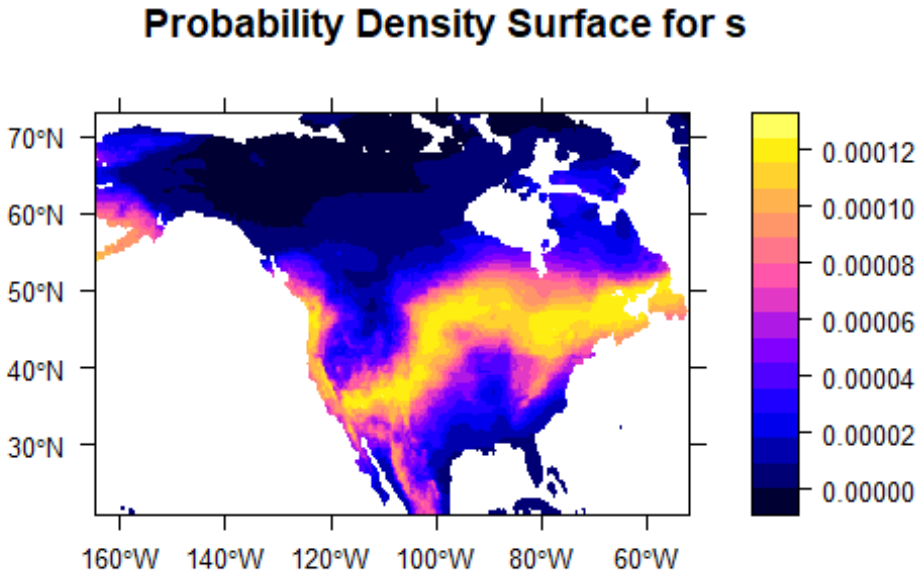

Probability Density Surface for p
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
1e-04
8e-05
6e-05
4e-05
2e-05
0e+00

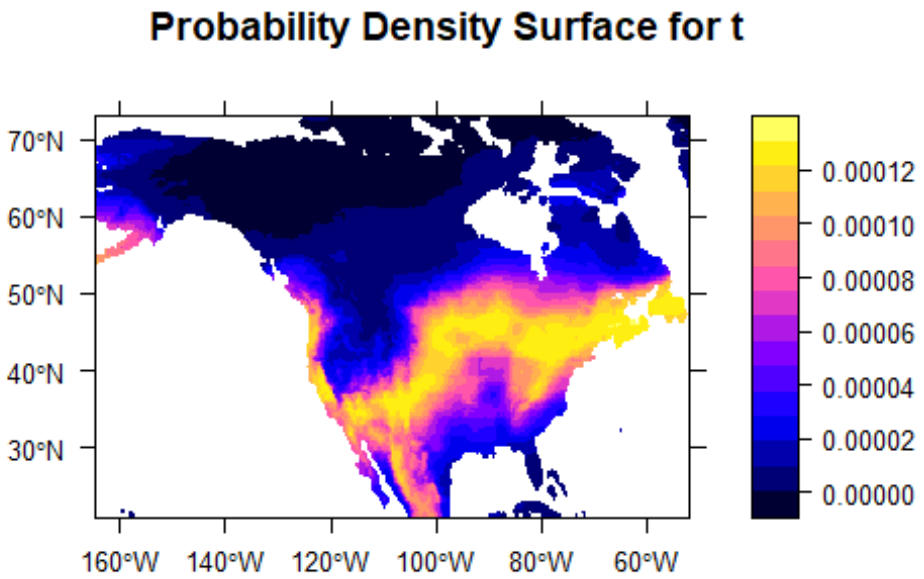
Probability Density Surface for q

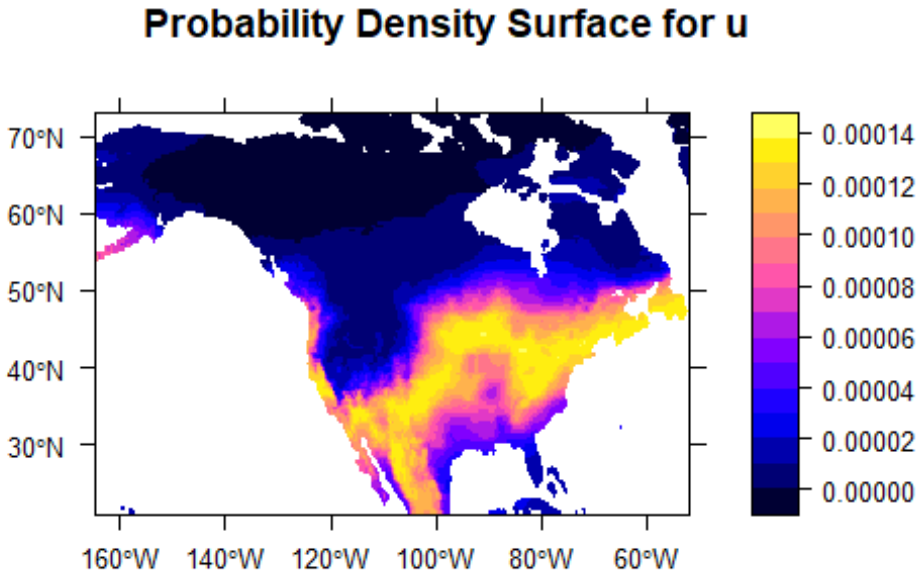


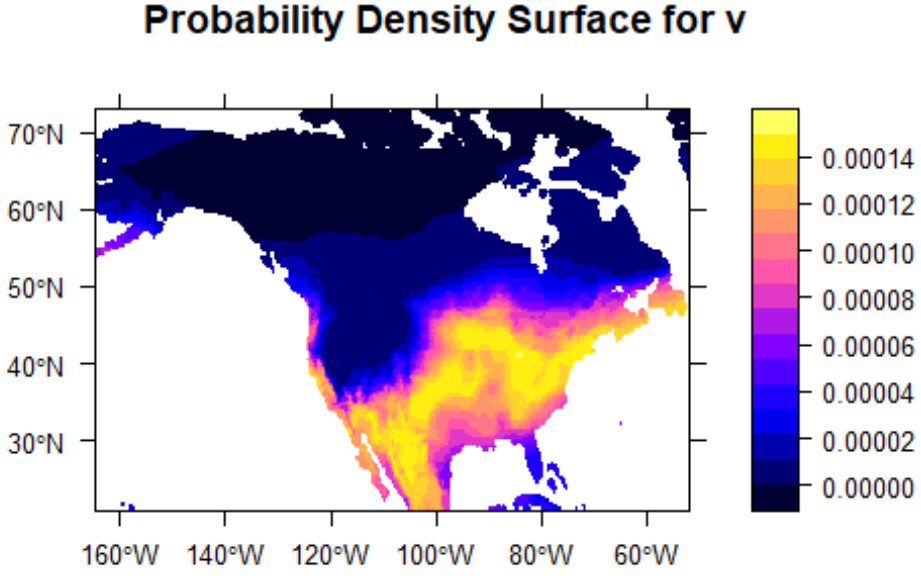
Probability Density Surface for r

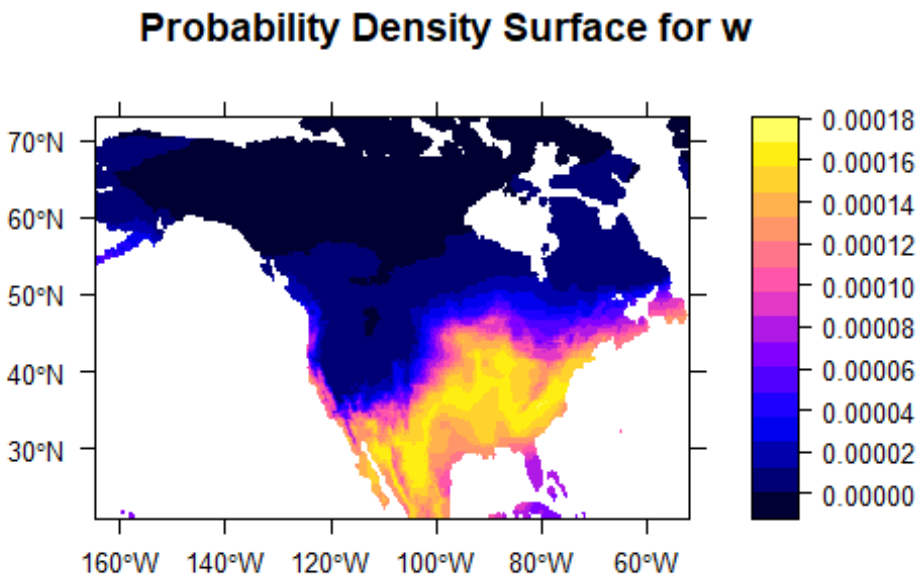


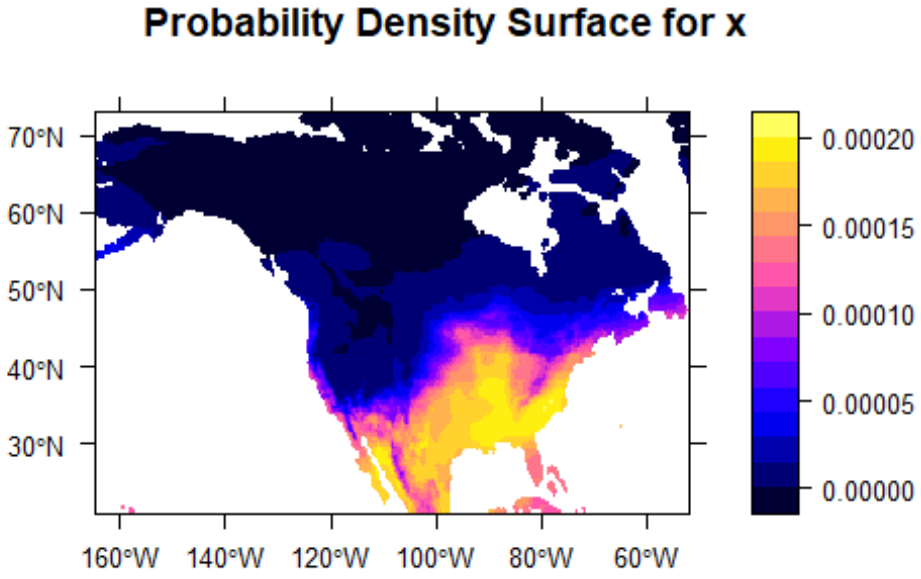

Probability Density Surface for s
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000

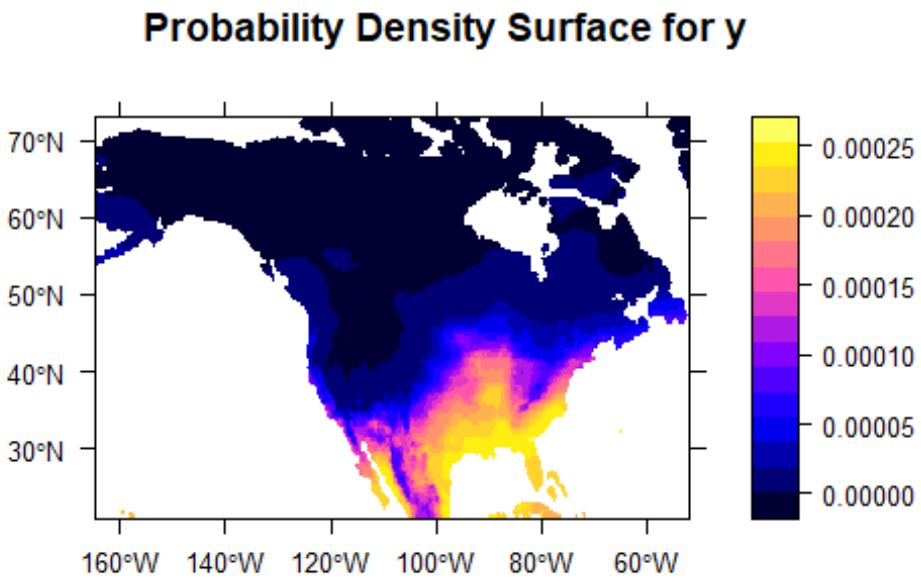

Probability Density Surface for t
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000

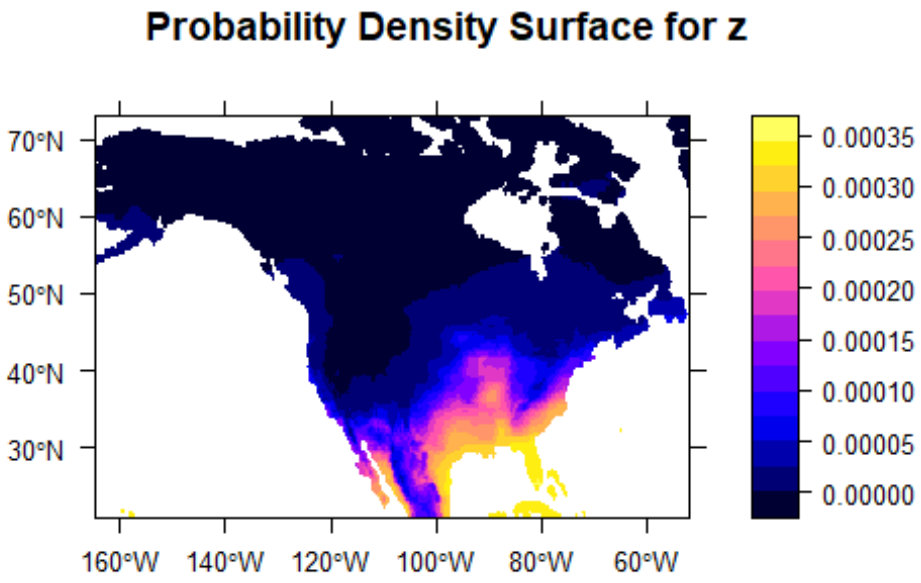

Probability Density Surface for u
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00014
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000


Probability Density Surface for v
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00014
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000


Probability Density Surface for w
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00018
0.00016
0.00014
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000


Probability Density Surface for x
70°N
60°N
50°N
40°N
30°N
160°W
140°W
120°W
100°W
80°W
60°W
0.00020
0.00015
0.00010
0.00005
0.00000

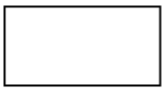




Create SpatialPolygons with two polygons

```
p1 <- c(-100,60,-100,65,-110,65,-110,60,-100,60)
p1 <- matrix(p1, 5,2, byrow = T)
p1 <- Polygon(p1)
```

```
p1 <- Polygons(list(p1), "p1")
p2 <- c(-100,40,-100,45,-110,45,-110,40,-100,40)
p2 <- matrix(p2, 5,2, byrow = T)
p2 <- Polygon(p2)
p2 <- Polygons(list(p2), "p2")
p12 <- SpatialPolygons(list(p1,p2),1:2)
plot(p12)
```



Create data.frame with two points

```
pp1 <- c(-100,45)
pp2 <- c(-100,60)
pp12 <- as.data.frame(rbind(pp1,pp2))
```

Caculate odds ratio for the two polygons created above

```
oddsRatio(asn, p12)
```

```
## $`P1/P2_odds_ratio`
##            a            b            c            d            e
## 4.307519e+03 2.638083e+03 1.526604e+03 8.436650e+02 4.495768e+02
##            f            g            h            i            j
## 2.327949e+02 1.177681e+02 5.838617e+01 2.839553e+01 1.353854e+01
##            k            l            m            n            o
## 6.318598e+00 2.881820e+00 1.282147e+00 5.550810e-01 2.330232e-01
##            p            q            r            s            t
## 9.449897e-02 3.692286e-02 1.389081e-02 5.037257e-03 1.764351e-03
##            u            v            w            x            y
```

```
## 5.985062e-04 1.972676e-04 6.339322e-05 1.992382e-05 6.138387e-06
##            z
## 1.856493e-06
##
## $`ratio of numbers of cells in two polygons`
## [1] 1
```

Caculate odds ratio for the two points created above

```
oddsRatio(asn, pp12)
```

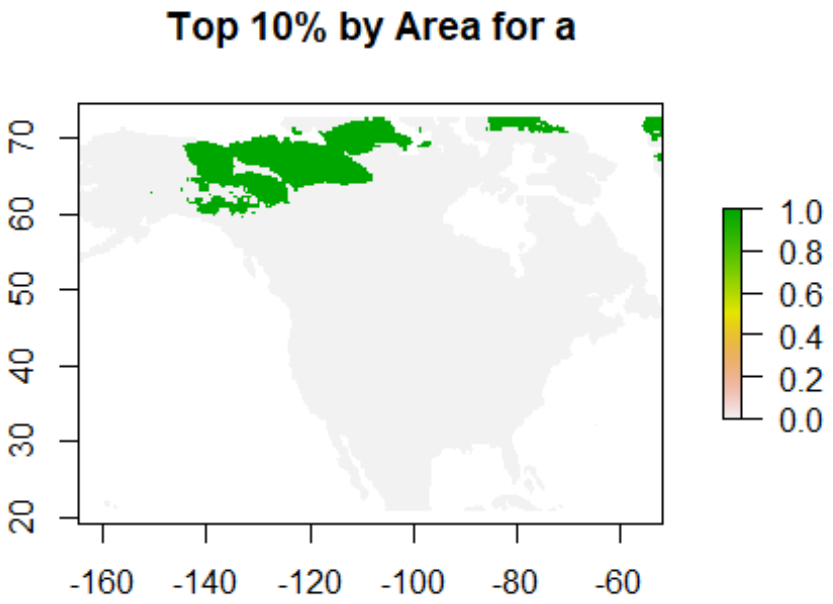
```
## $`P1/P2_odds_ratio`
##            a            b            c            d            e
## 1.244449e-08 4.779468e-08 1.808158e-07 6.738255e-07 2.473504e-06
##            f            g            h            i            j
## 8.943996e-06 3.185695e-05 1.117714e-04 3.862881e-04 1.315067e-03
##            k            l            m            n            o
## 4.410032e-03 1.456784e-02 4.740326e-02 1.519429e-01 4.797456e-01
##            p            q            r            s            t
## 1.492103e+00 4.571335e+00 1.379562e+01 4.101001e+01 1.200842e+02
##            u            v            w            x            y
## 3.463612e+02 9.840613e+02 2.754027e+03 7.592240e+03 2.061716e+04
##            z
## 5.514996e+04
##
## $`odds of a pixel to the odds of the max/min pixel`
##   ratioToMax.a ratioToMax.b ratioToMax.c ratioToMax.d ratioToMax.e
## 1 2.094192e-11 3.282734e-10  4.24455e-09 4.544198e-08 3.696248e-07
## 2 2.670406e-03 9.735931e-03  3.26940e-02 8.262258e-02 1.634092e-01
##   ratioToMax.f ratioToMax.g ratioToMax.h ratioToMax.i ratioToMax.j
## 1 2.309097e-06 1.221952e-05 5.694376e-05 0.0002362105 0.0008761164
## 2 2.641722e-01 3.786819e-01 4.892904e-01 0.5798734058 0.6326543312
##   ratioToMax.k ratioToMax.l ratioToMax.m ratioToMax.n ratioToMax.o
## 1  0.002915017  0.007795595   0.01491739  0.006763223   0.03541972
## 2  0.609662750  0.450544801   0.23110862  0.070630011   0.10465197
##   ratioToMax.p ratioToMax.q ratioToMax.r ratioToMax.s ratioToMax.t
## 1    0.1397912    0.4145120   0.84132231   1.18062565   1.27708736
## 2    0.1304856    0.1110991   0.06669533   0.02946187   0.01050098
##   ratioToMax.u ratioToMax.v ratioToMax.w ratioToMax.x ratioToMax.y
## 1  1.131160040 0.8524267526  0.565086007 3.398200e-01 1.700921e-01
## 2  0.003137033 0.0008215826  0.000194879 4.128849e-05 6.946884e-06
##   ratioToMax.z ratioToMin.a ratioToMin.b ratioToMin.c ratioToMin.d
## 1 6.421370e-02 7.169965e+10 1.047379e+09 1.791942e+07 3.671458e+05
## 2 8.551729e-07 2.646680e+17 1.112807e+15 5.601500e+12 3.514137e+10
##   ratioToMin.e ratioToMin.f ratioToMin.g ratioToMin.h ratioToMin.i
## 1      9559.58     350.9178     20.26651     1.911242 9.966561e-01
## 2 298570754.84 3863909.9809  82600.34476  2935.584705 1.091443e+04
##   ratioToMin.j ratioToMin.k ratioToMin.l ratioToMin.m ratioToMin.n
## 1     83.60916     11698.49      3014937   1040393968 2.315550e+19
## 2 362976.90841  22042755.32   2169295297 273195801360 7.000247e+18
##   ratioToMin.o ratioToMin.p ratioToMin.q ratioToMin.r ratioToMin.s
```

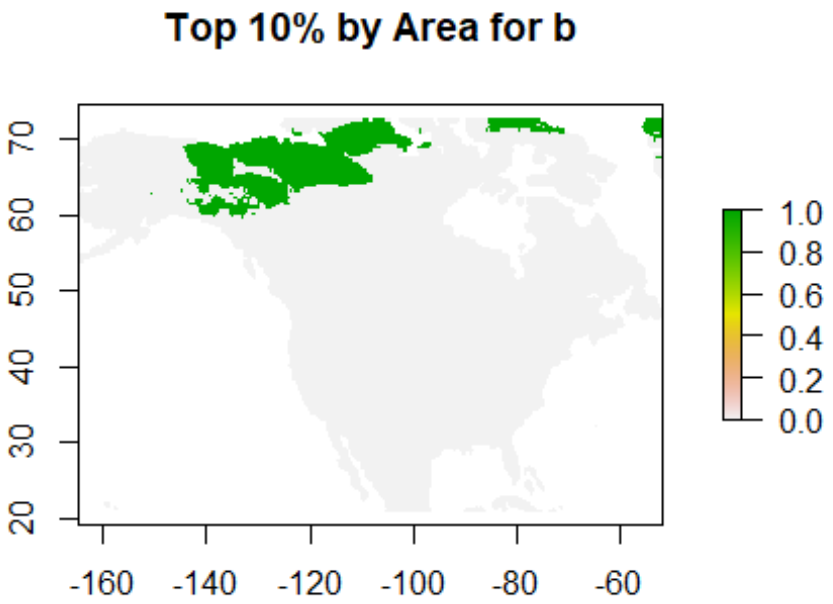
```
## 1 1.130091e+17 5.901633e+14 3.349025e+12  21759061135  179421903.5
## 2 1.196162e+16 2.235624e+13 4.725311e+10    121861383     430923.5
##   ratioToMin.t ratioToMin.u ratioToMin.v ratioToMin.w ratioToMin.x
## 1  2118094.797  37965.88735   3596.69144   53927.0407  1363758.896
## 2     2290.536     18.82119     15.46391     111.8092     1492.812
##   ratioToMin.y ratioToMin.z
## 1  65782920.70   4478501418
## 2     33448.05      1010908
```

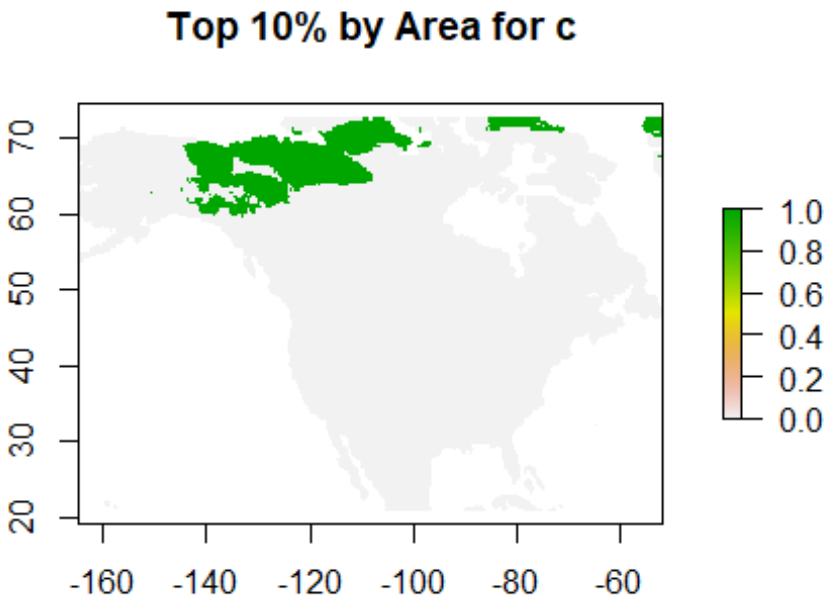
Binary reclassification

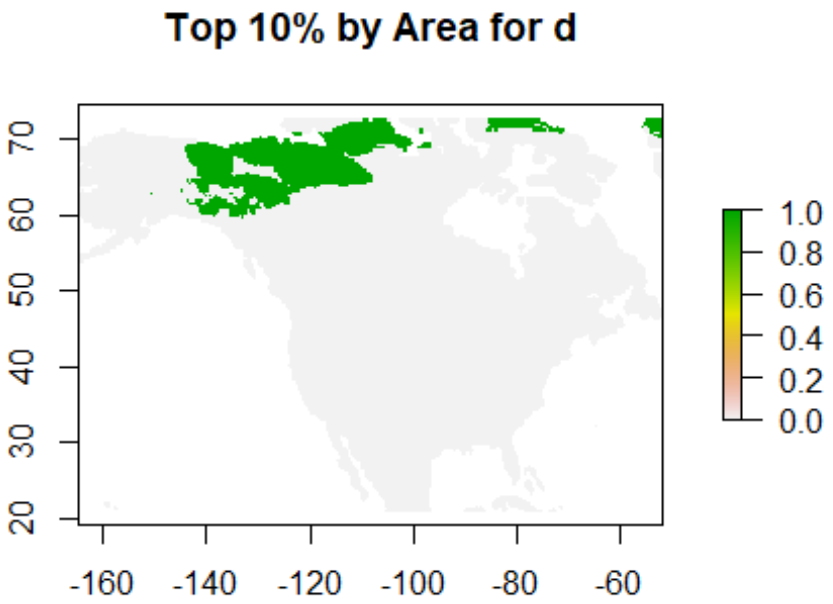
Top 10% of probability surface (defined by % area)

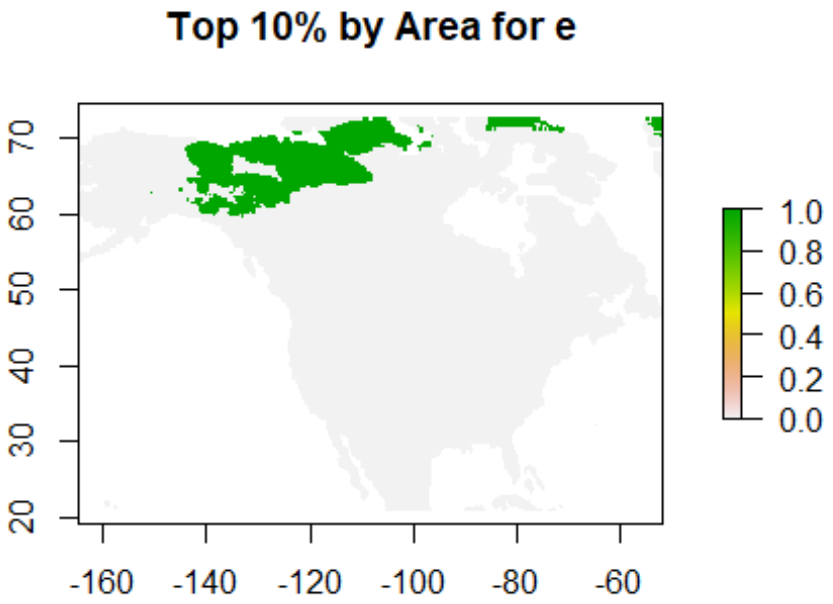
```
qtlRaster(asn, threshold = 0.1, thresholdType = 2)
```

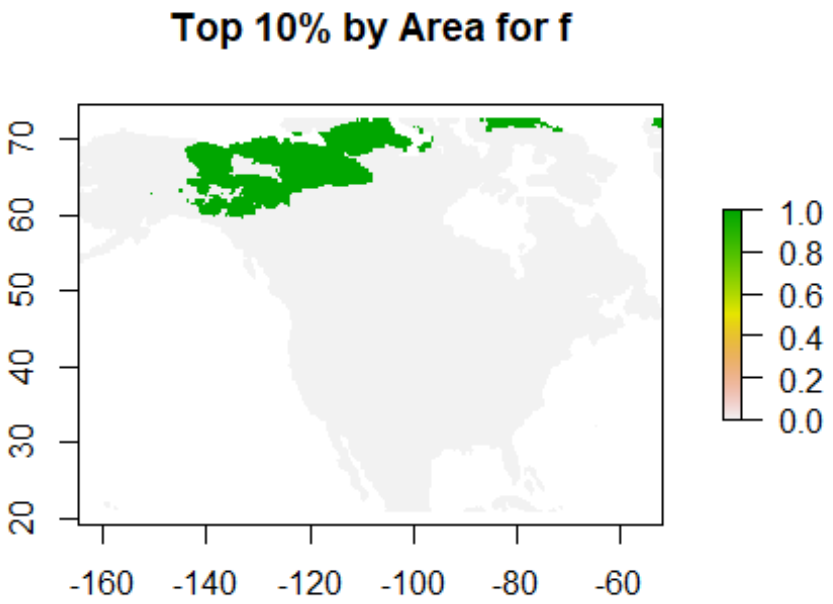

Top 10% by Area for a
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

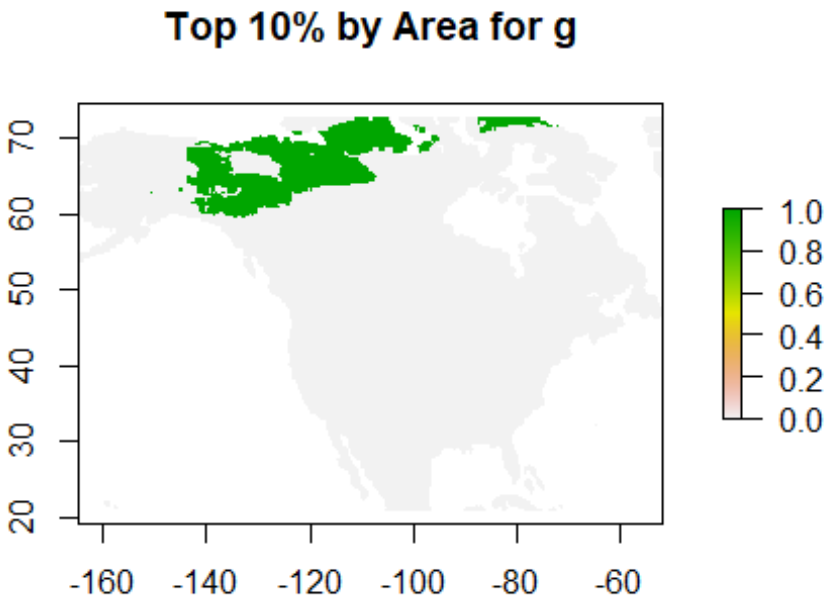

Top 10% by Area for b
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

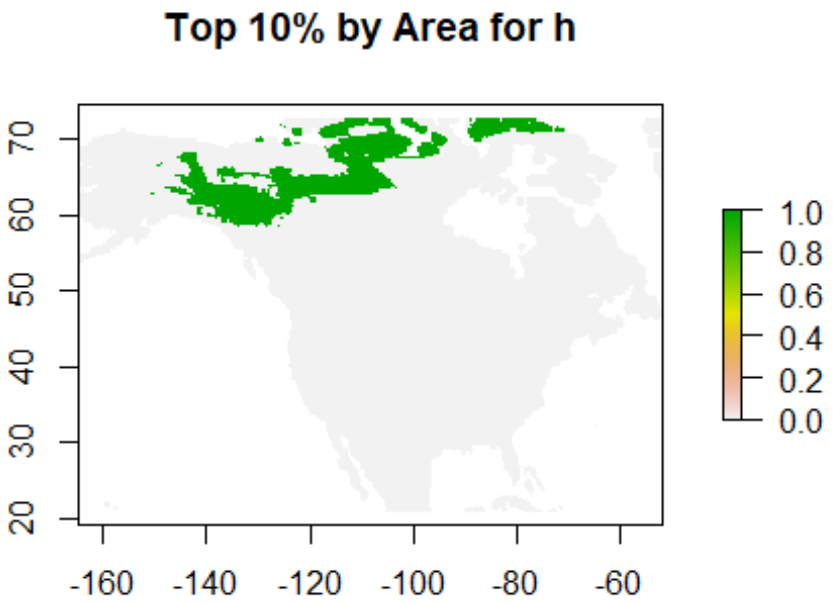

Top 10% by Area for c
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

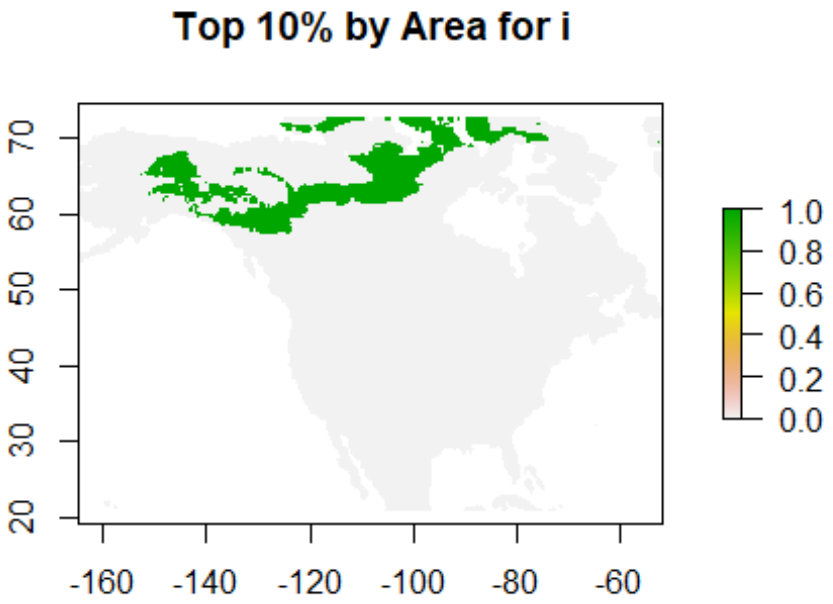

Top 10% by Area for d
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

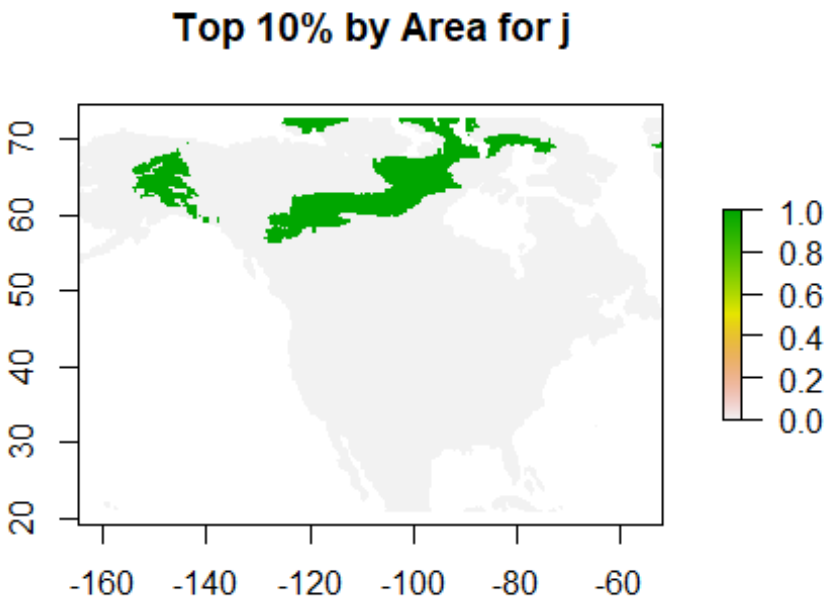

Top 10% by Area for e
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

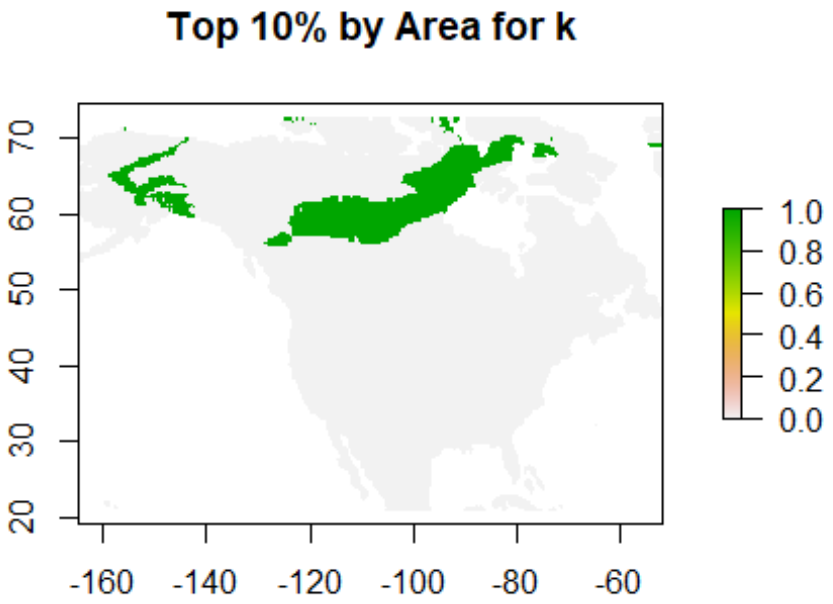

Top 10% by Area for f
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

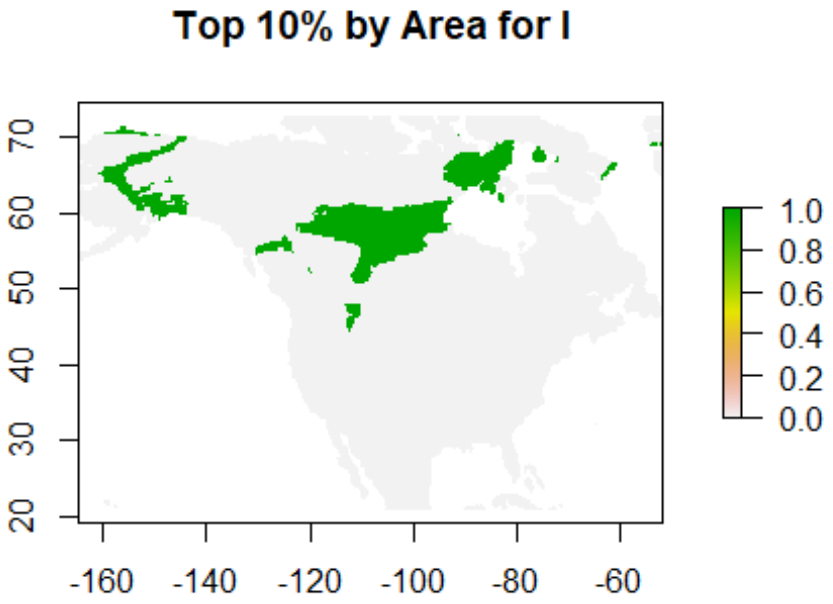

Top 10% by Area for g
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

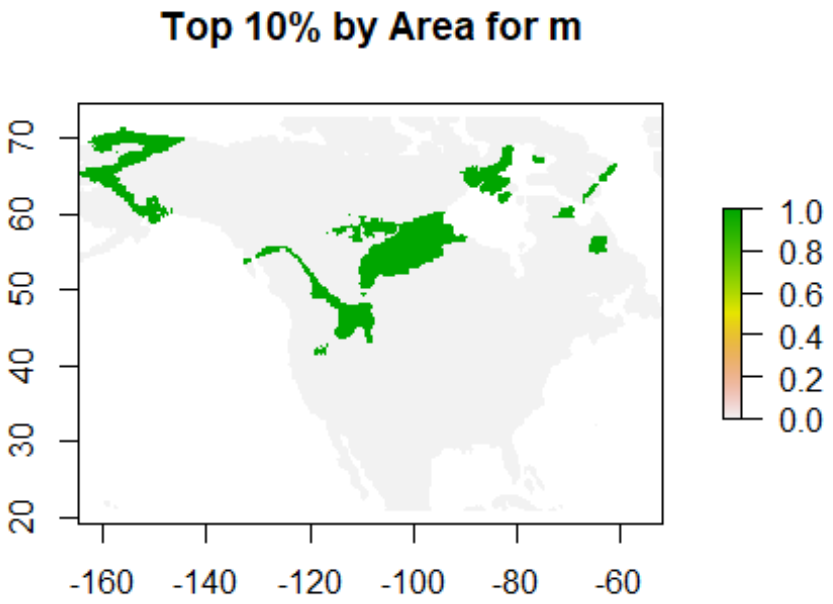

Top 10% by Area for h
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

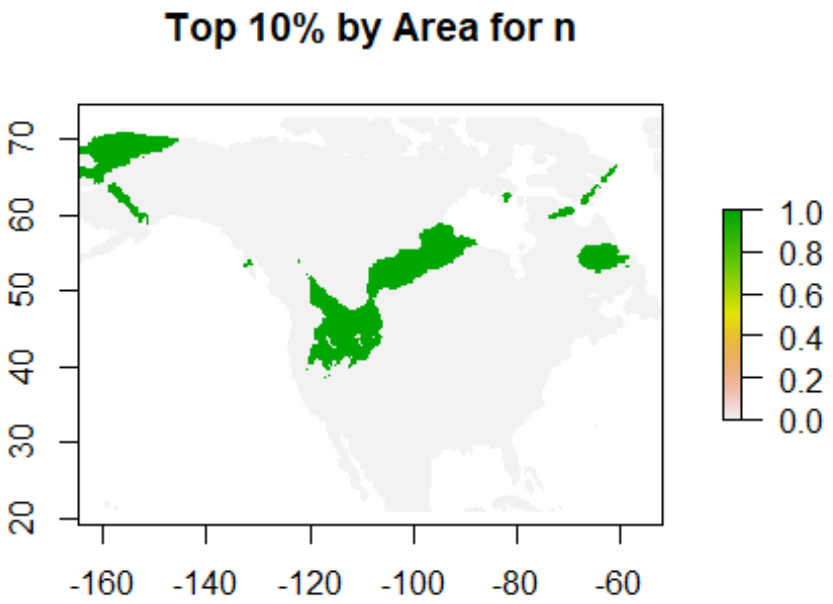

Top 10% by Area for i
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

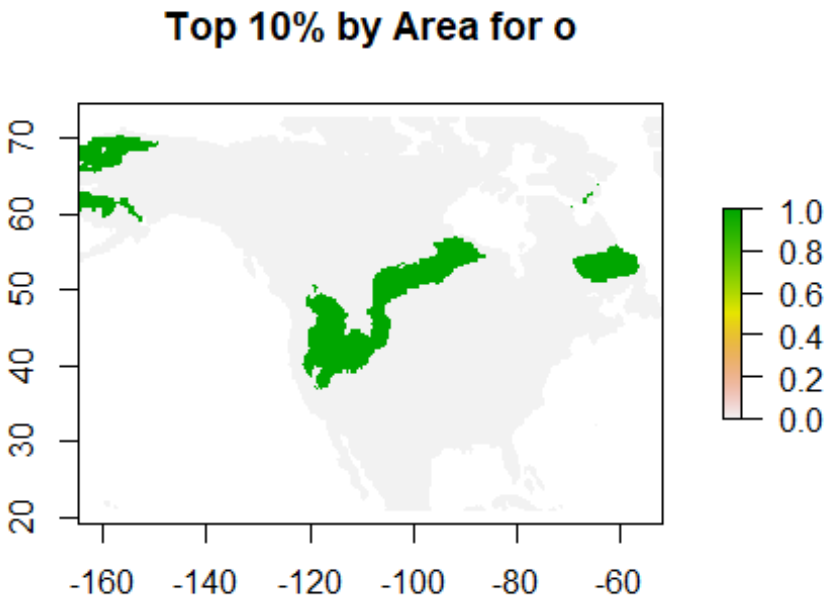

Top 10% by Area for j
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

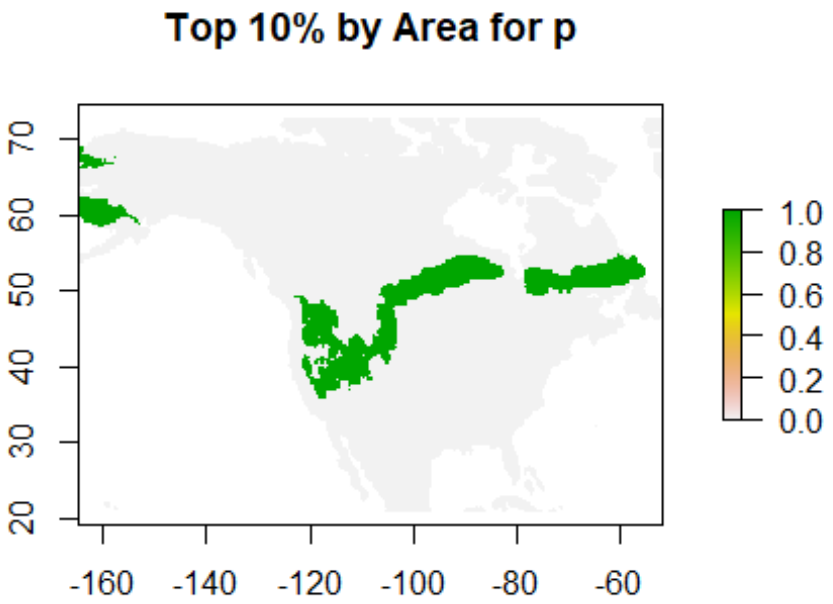

Top 10% by Area for k
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

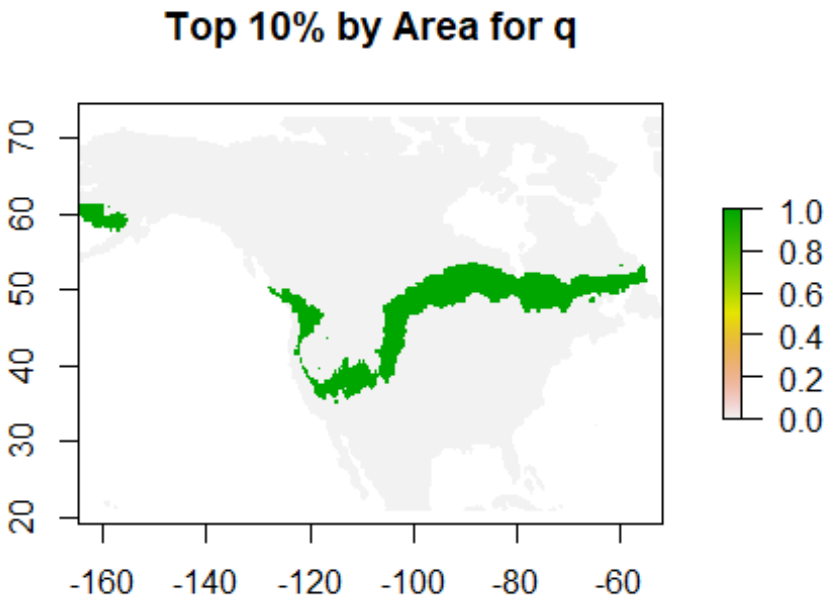

Top 10% by Area for l
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

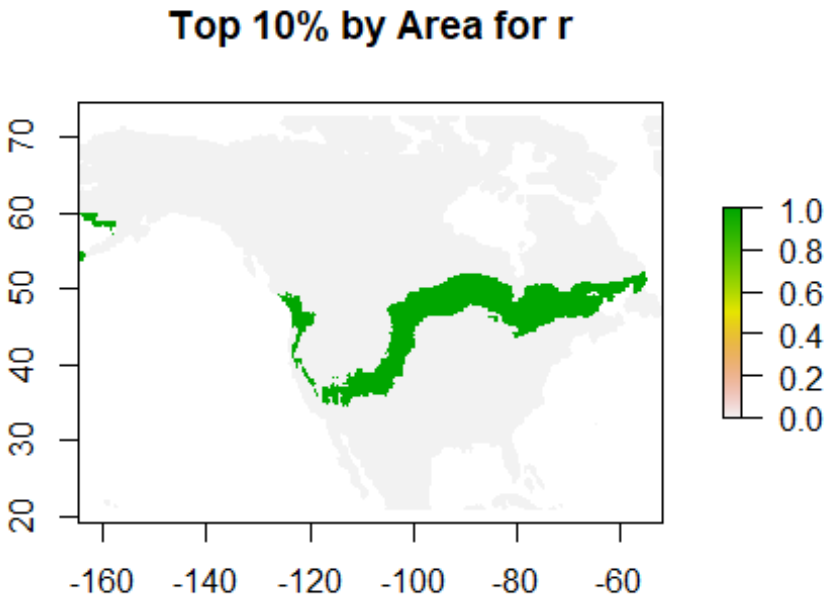

Top 10% by Area for m
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

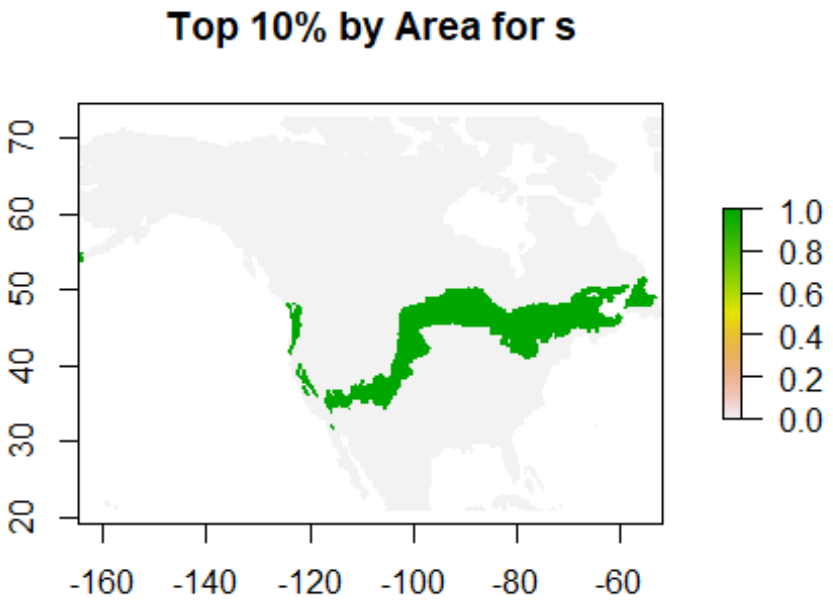

Top 10% by Area for n
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

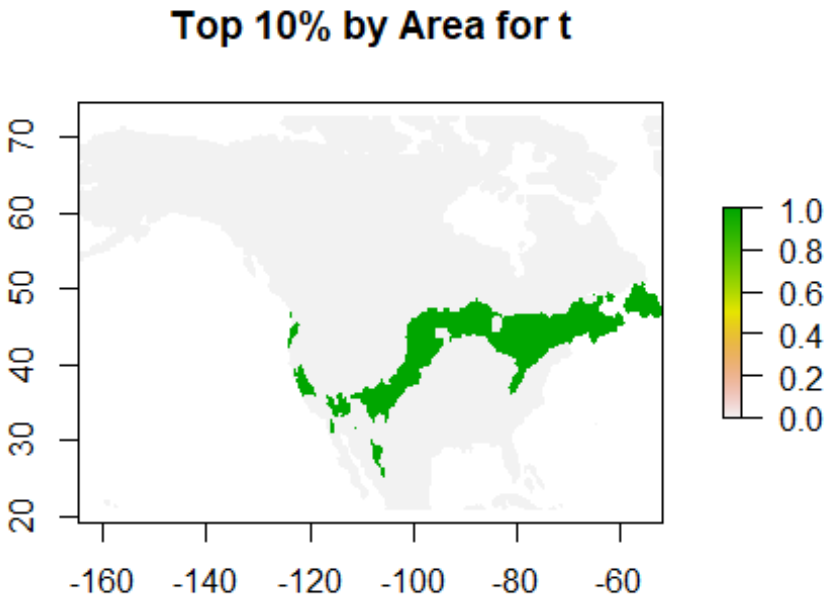

Top 10% by Area for o
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

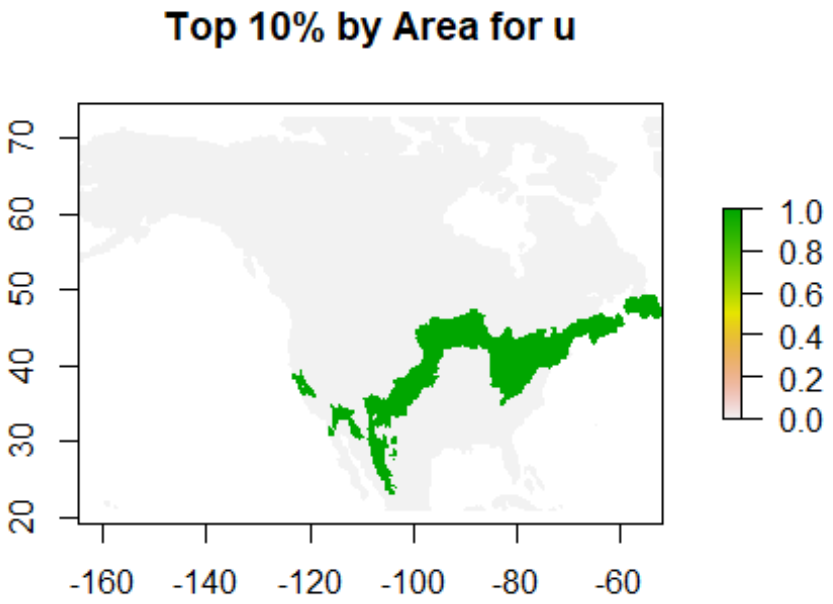

Top 10% by Area for p
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

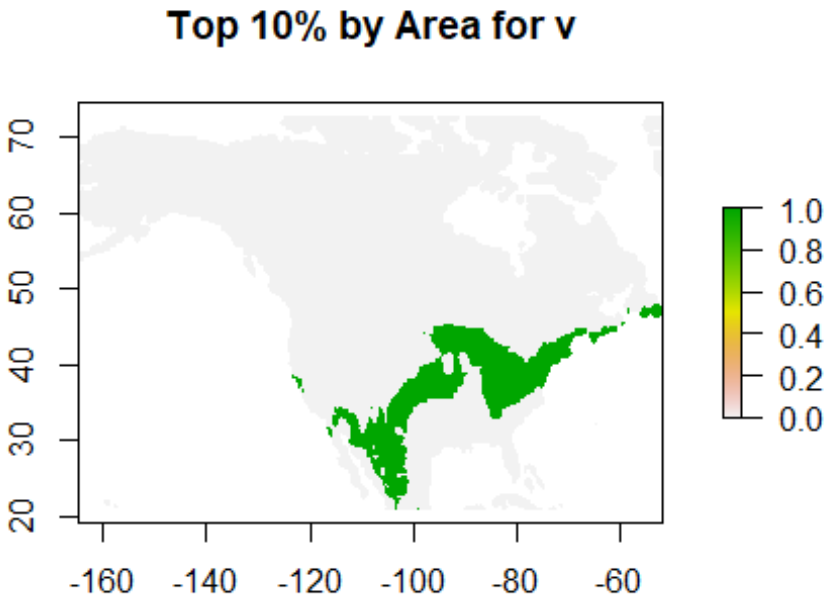

Top 10% by Area for q
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

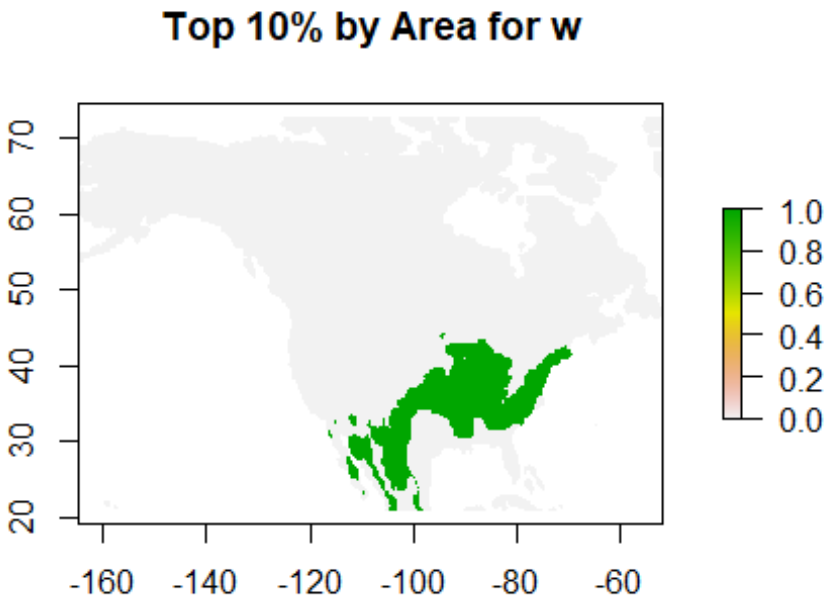

Top 10% by Area for r
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

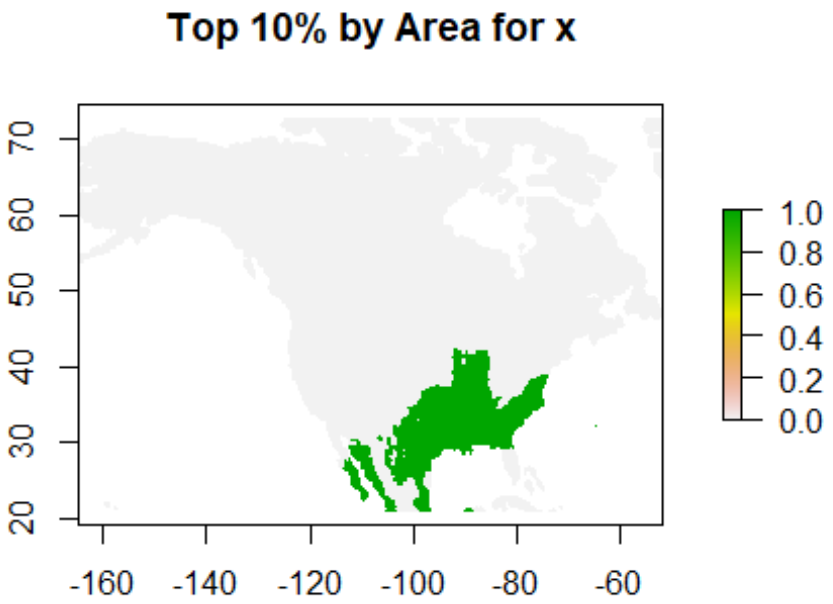

Top 10% by Area for s
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

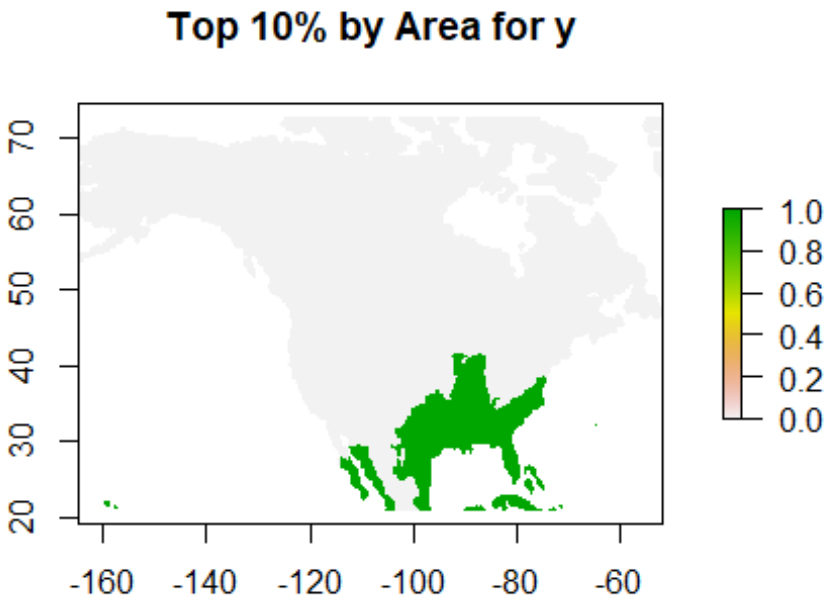

Top 10% by Area for t
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0

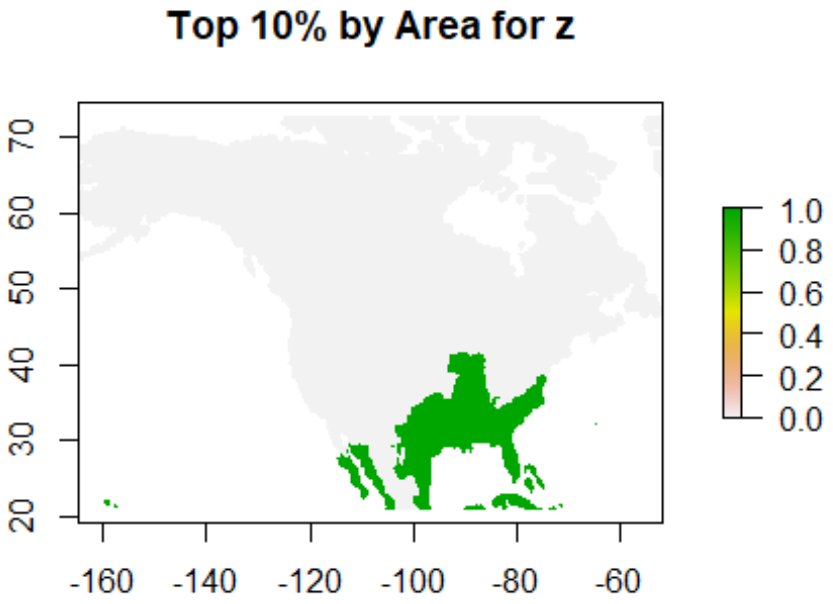

Top 10% by Area for u
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0


Top 10% by Area for v
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0


Top 10% by Area for w
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0


Top 10% by Area for x
70
60
50
40
30
20
-160
-140
-120
-100
-80
-60
1.0
0.8
0.6
0.4
0.2
0.0





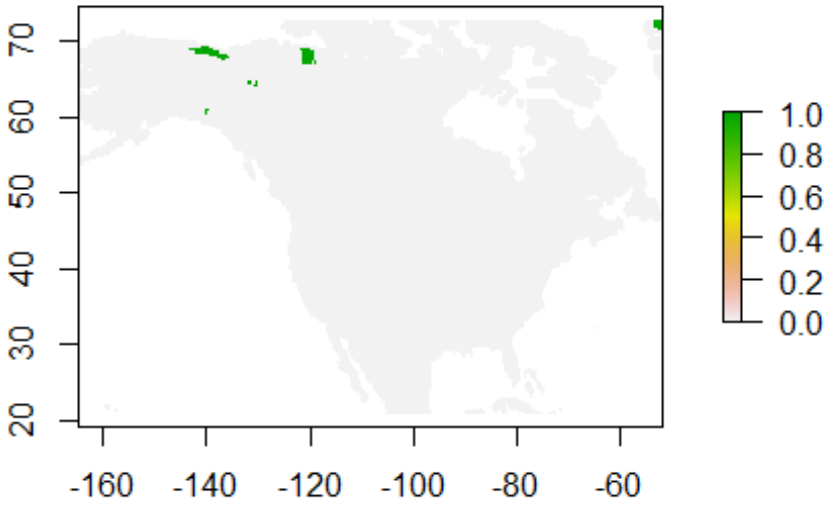

```
## class       : RasterStack
## dimensions  : 156, 339, 52884, 26  (nrow, ncol, ncell, nlayers)
## resolution  : 0.333, 0.333  (x, y)
## extent      : -164.682, -51.795, 20.97926, 72.92726  (xmin, xmax, ymin,
```

```
ymax)
## coord. ref. : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84
+towgs84=0,0,0
## names       : a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, ...
## min values  : 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
## max values  : 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
```

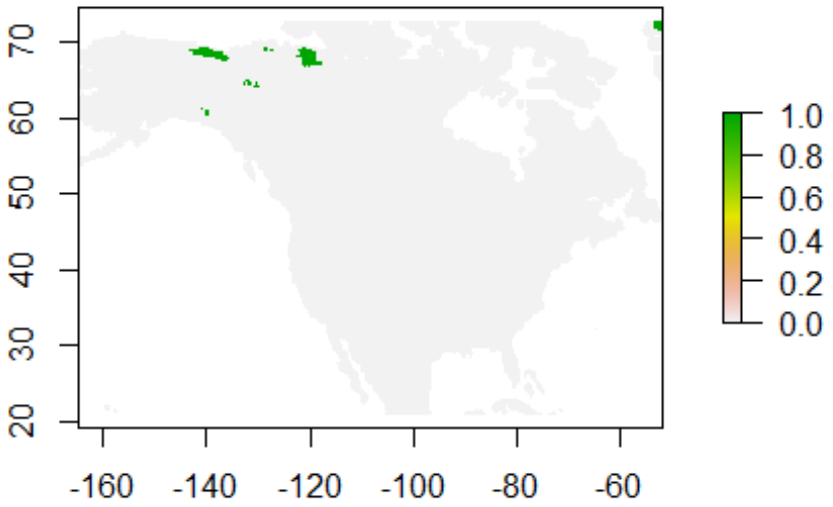
Top 10% of probability surface (defined by % cumulative probability)

```
qtlRaster(asn, threshold = 0.1, thresholdType = 1)
```

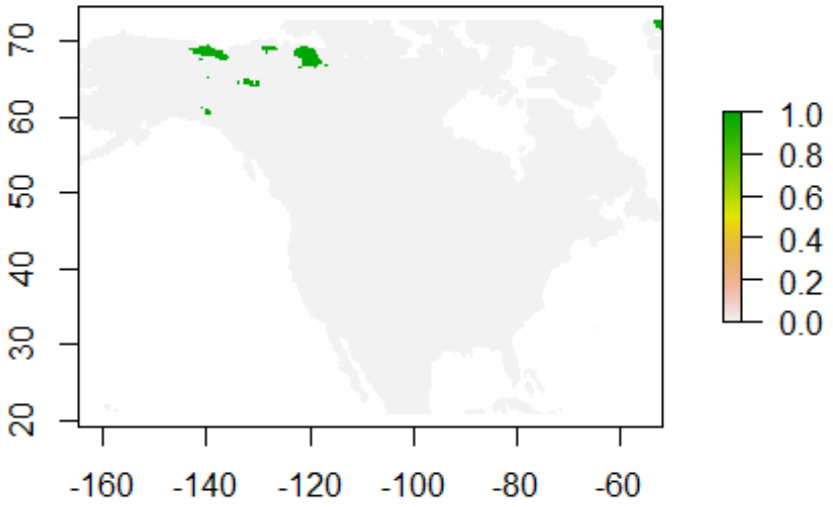
Top 10% by Cumulative Probability for a



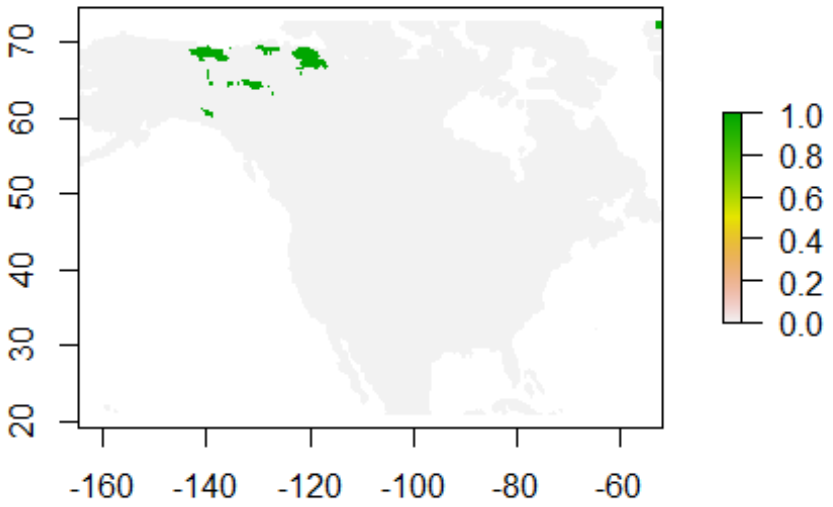
Top 10% by Cumulative Probability for b



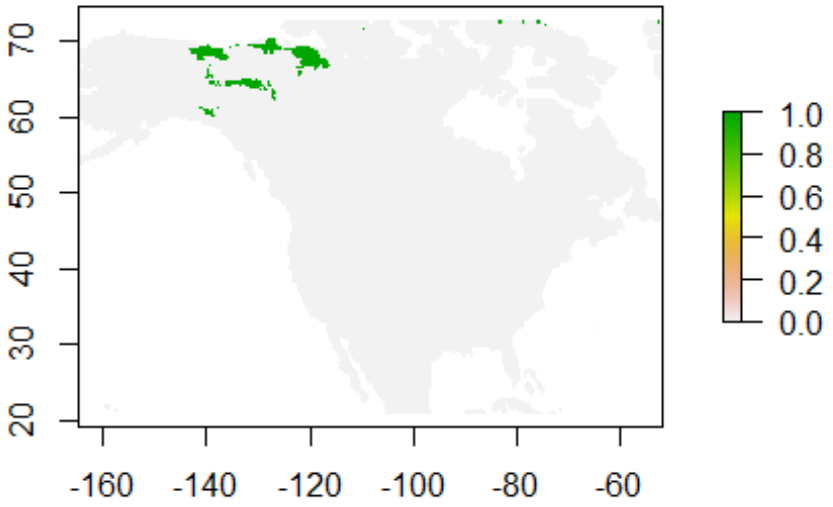
Top 10% by Cumulative Probability for c



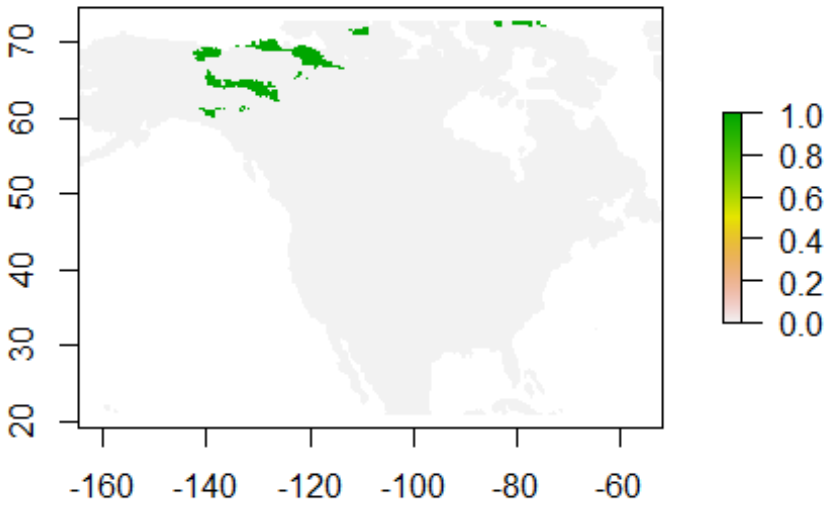
Top 10% by Cumulative Probability for d



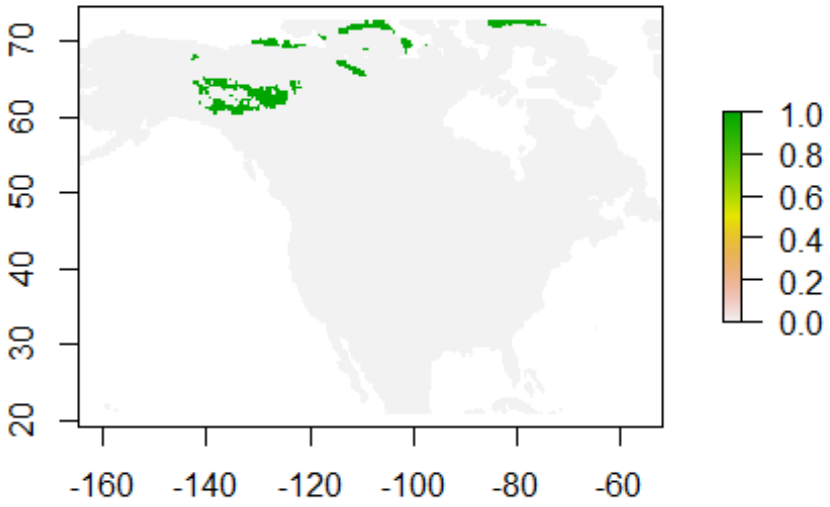
Top 10% by Cumulative Probability for e



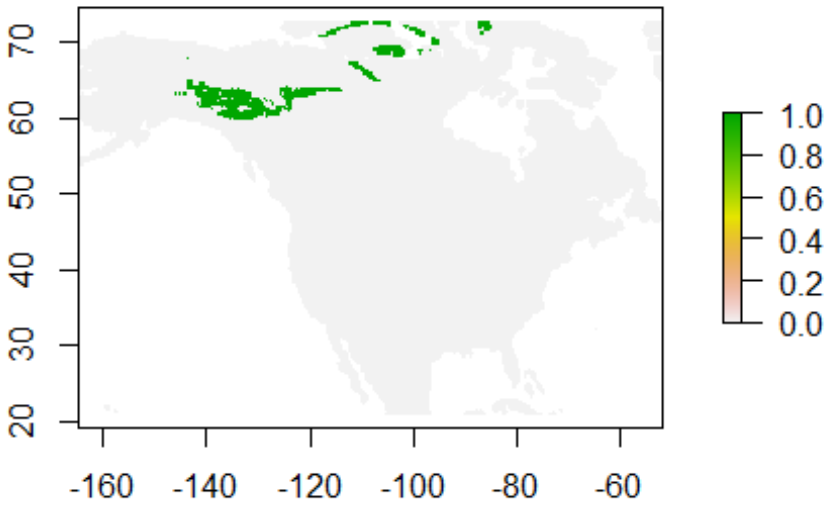
Top 10% by Cumulative Probability for f



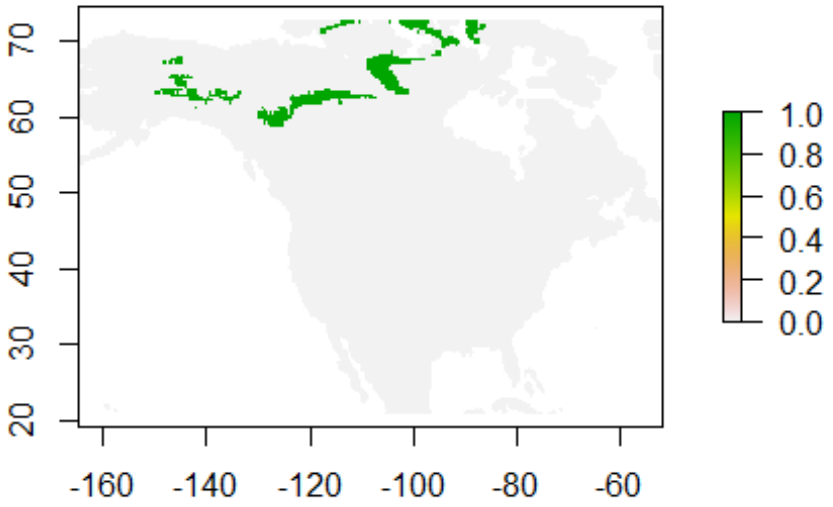
Top 10% by Cumulative Probability for g



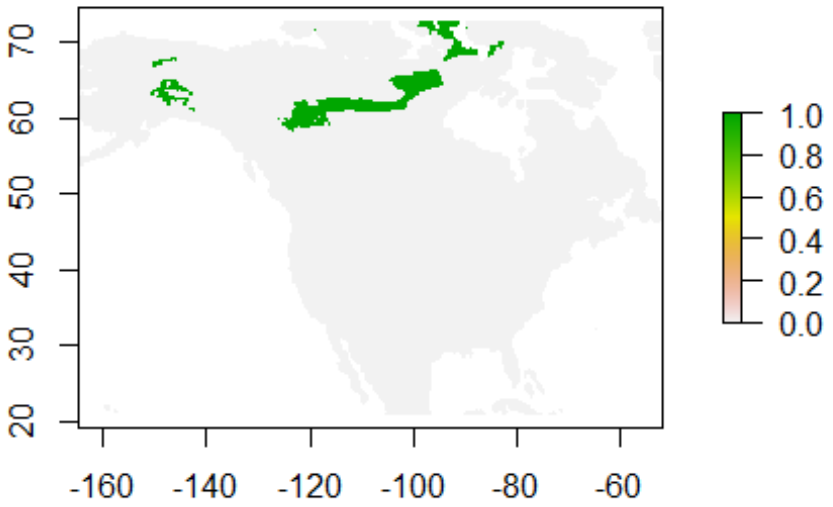
Top 10% by Cumulative Probability for h



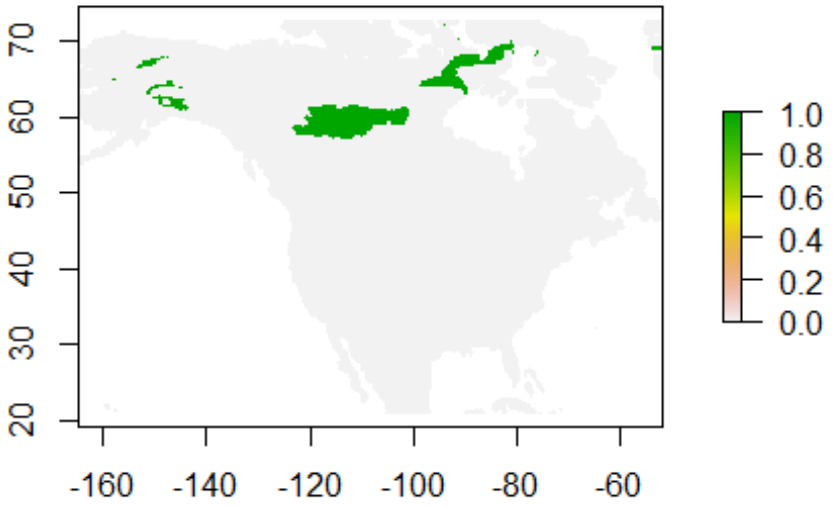
Top 10% by Cumulative Probability for i



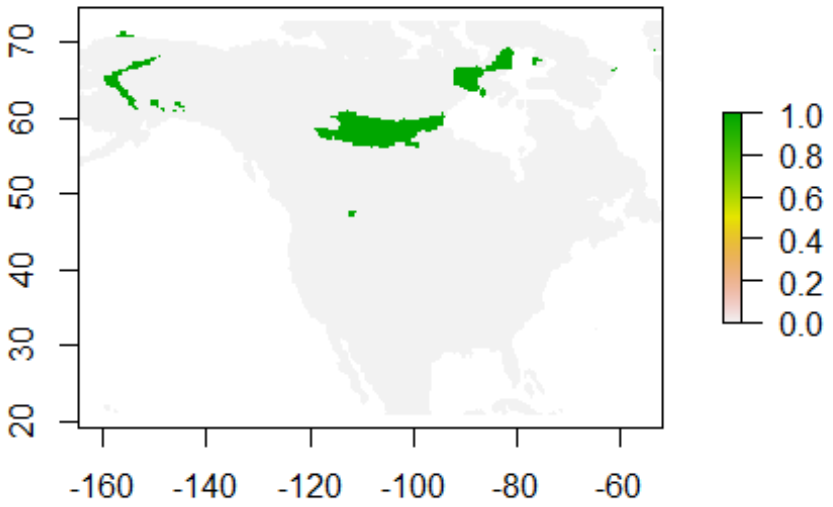
Top 10% by Cumulative Probability for j



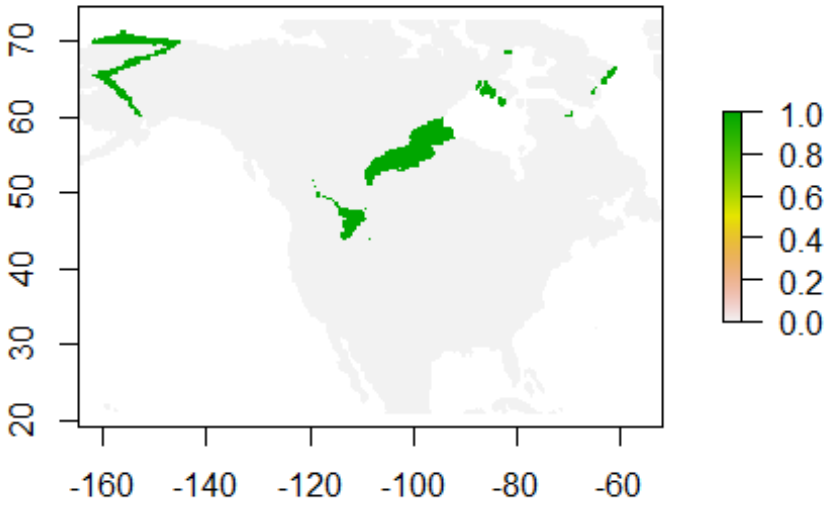
Top 10% by Cumulative Probability for k



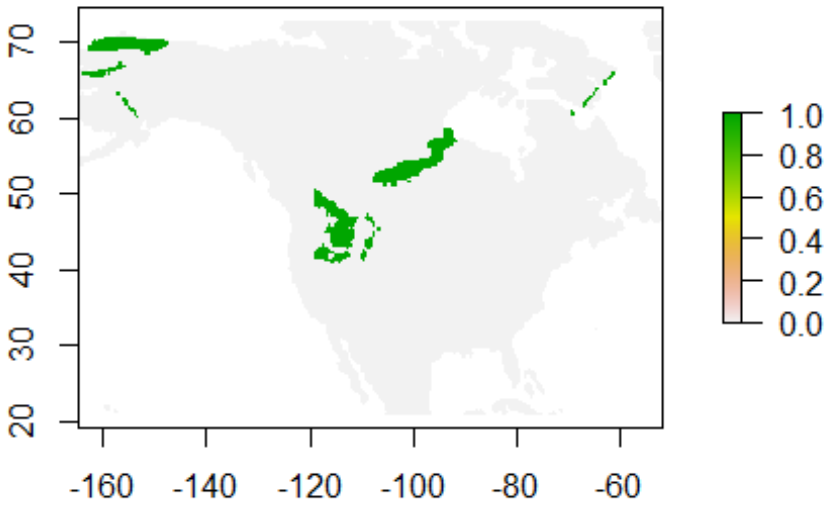
Top 10% by Cumulative Probability for l



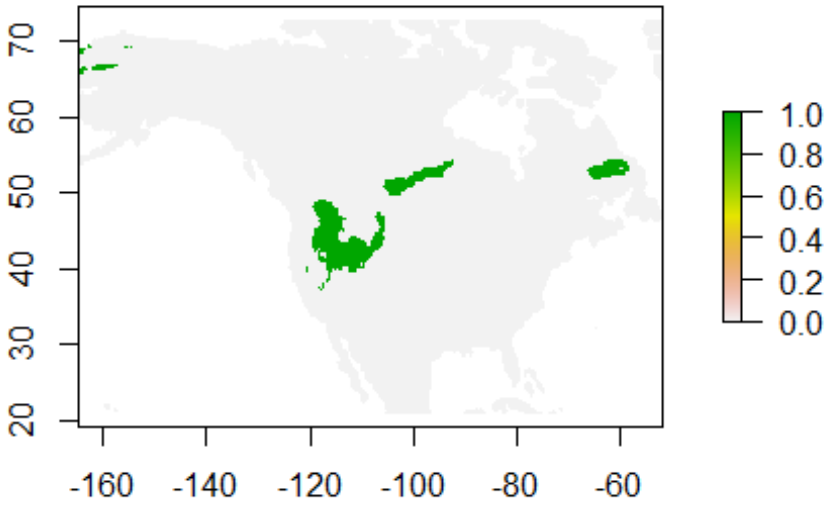
Top 10% by Cumulative Probability for m



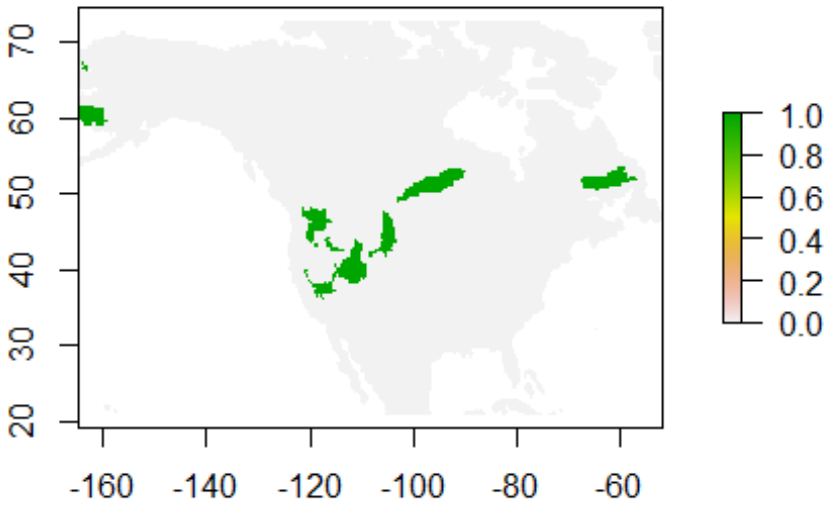
Top 10% by Cumulative Probability for n



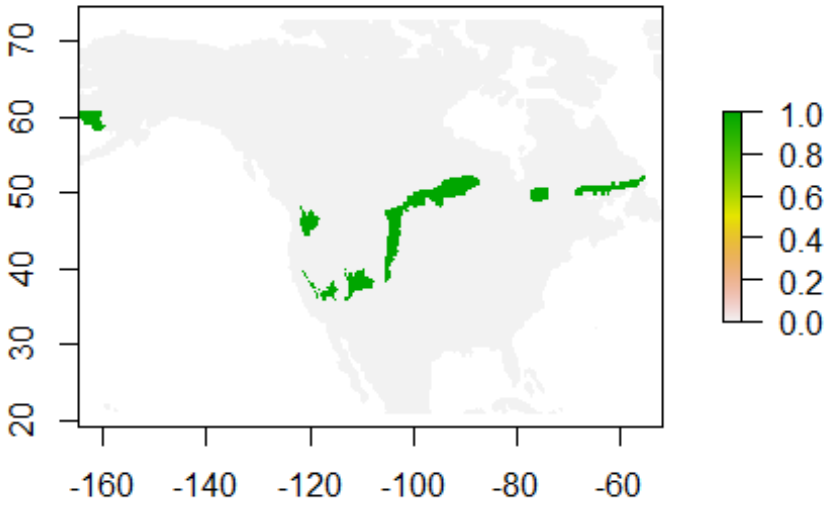
Top 10% by Cumulative Probability for o



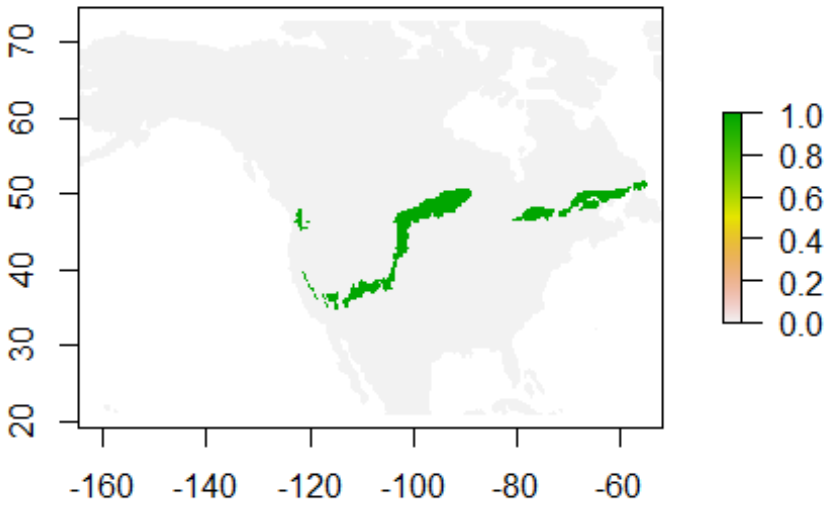
Top 10% by Cumulative Probability for p



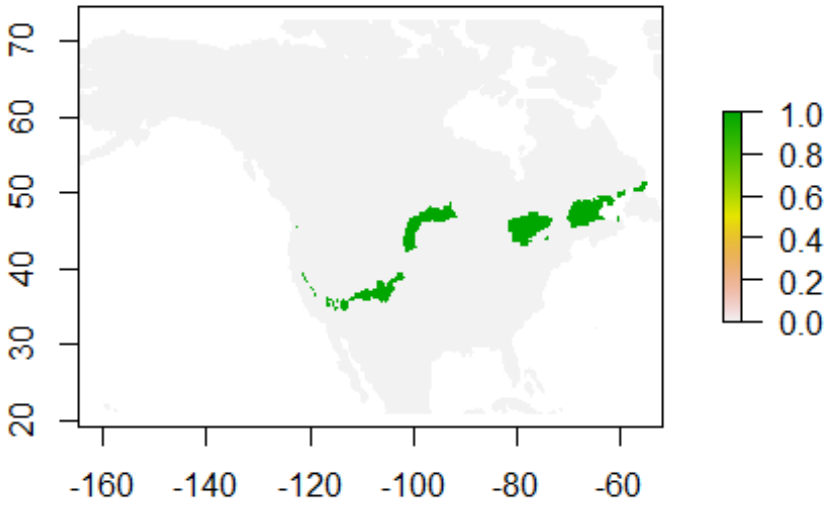
Top 10% by Cumulative Probability for q



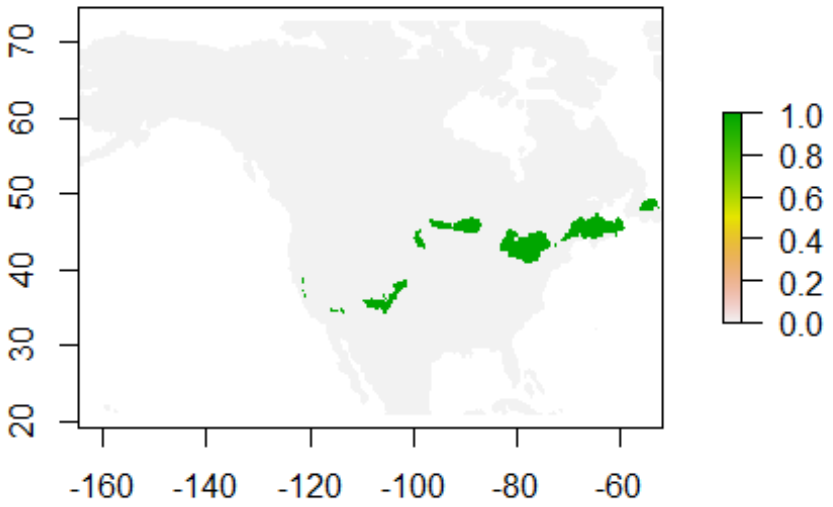
Top 10% by Cumulative Probability for r



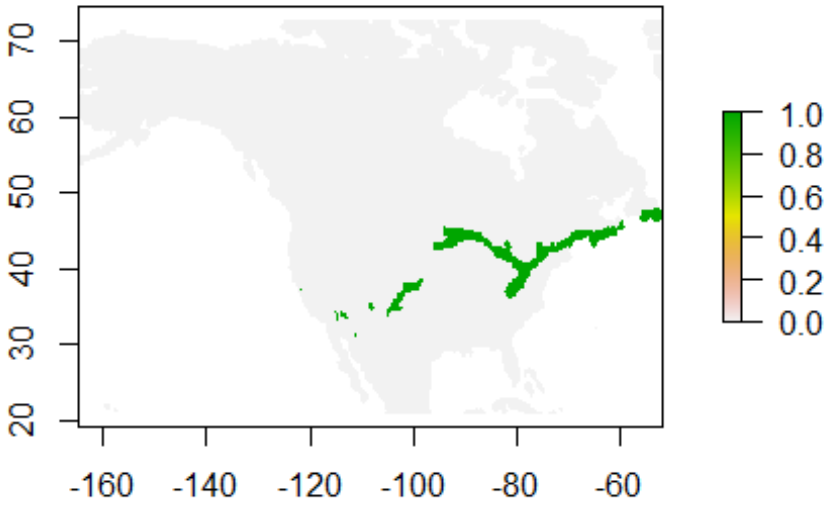
Top 10% by Cumulative Probability for s



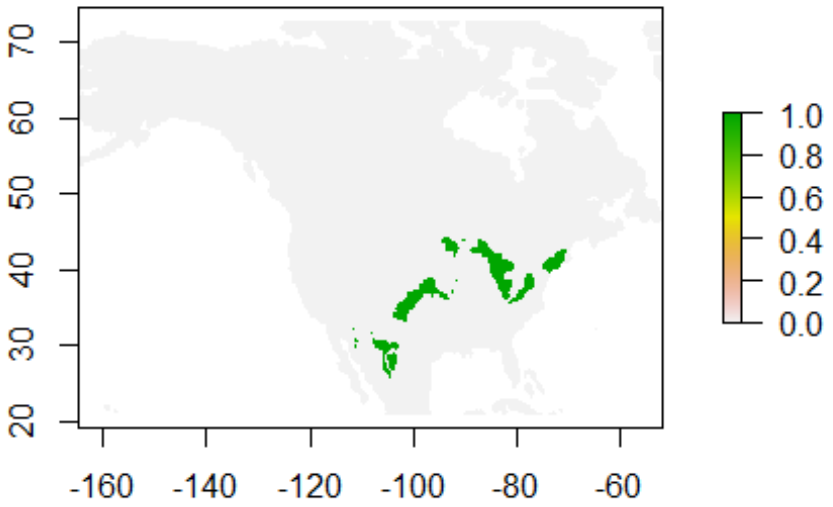
Top 10% by Cumulative Probability for t



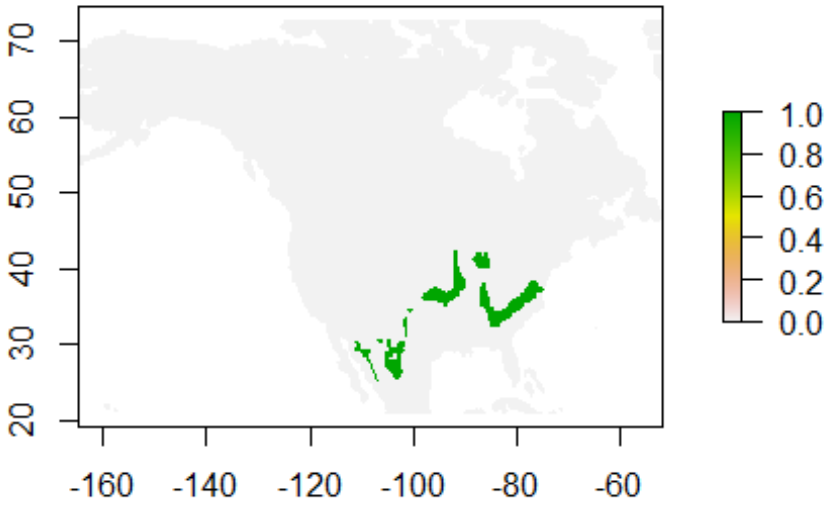
Top 10% by Cumulative Probability for u



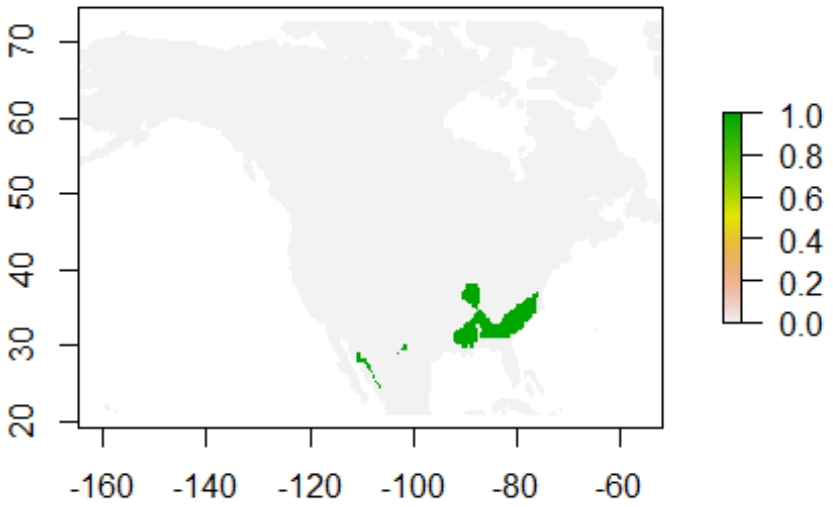
Top 10% by Cumulative Probability for v

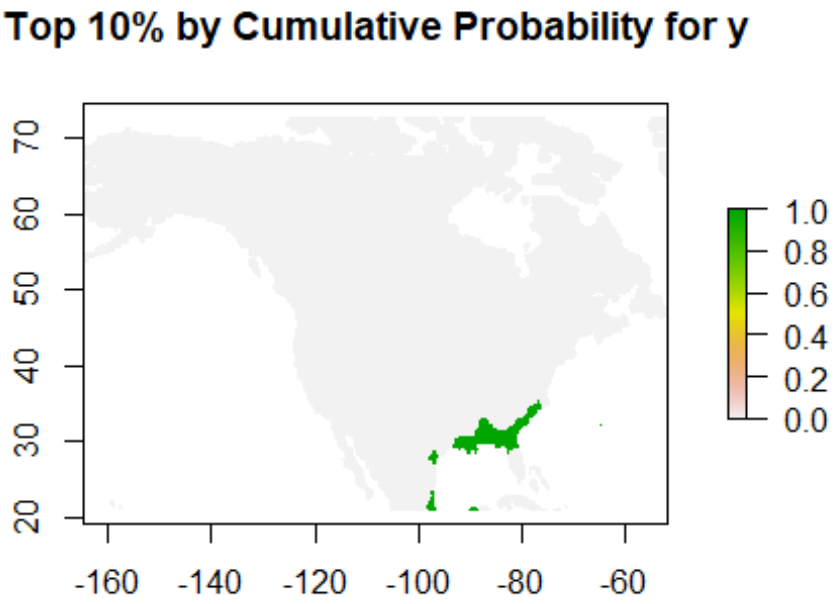


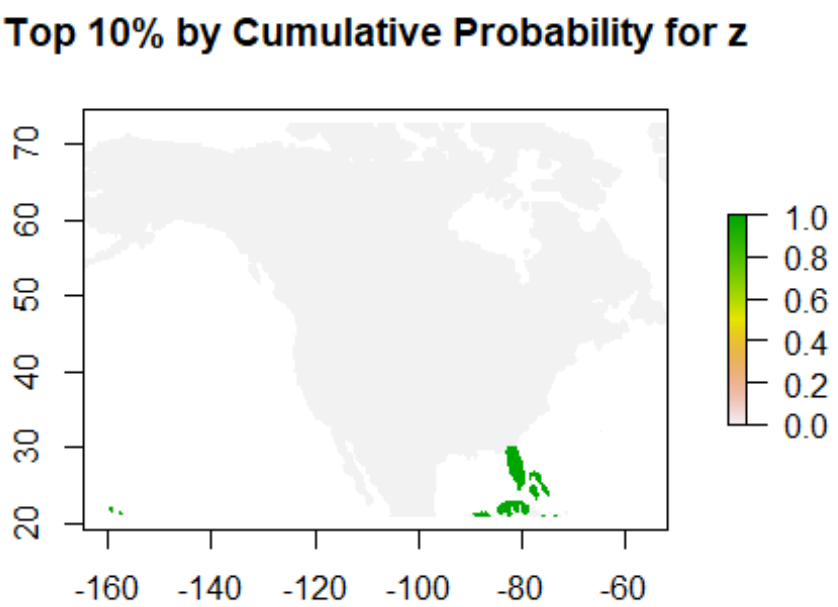
Top 10% by Cumulative Probability for w



Top 10% by Cumulative Probability for x





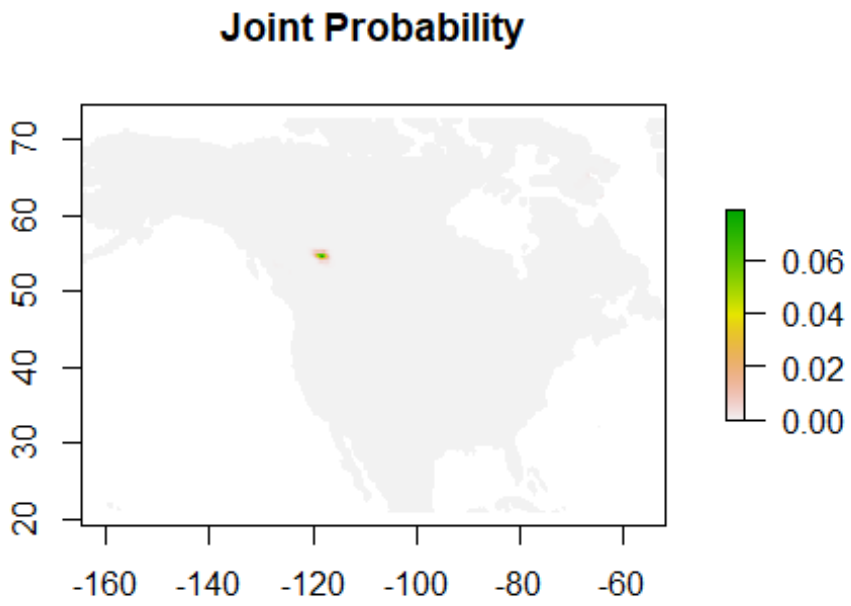



```
## class       : RasterStack
## dimensions  : 156, 339, 52884, 26  (nrow, ncol, ncell, nlayers)
## resolution  : 0.333, 0.333  (x, y)
## extent      : -164.682, -51.795, 20.97926, 72.92726  (xmin, xmax, ymin,
```

```
ymax)
## coord. ref. : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84
+towgs84=0,0,0
## names       : a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, ...
## min values  : 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
## max values  : 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
```

Joint probability for individuals of common origin

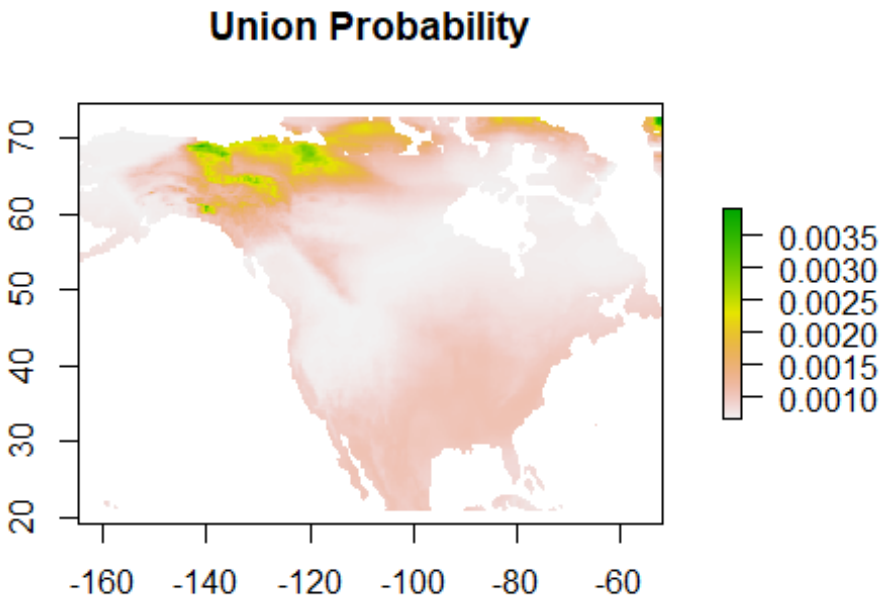
```
jointP(asn)
```




```
## class       : RasterLayer
## dimensions  : 156, 339, 52884  (nrow, ncol, ncell)
## resolution  : 0.333, 0.333  (x, y)
## extent      : -164.682, -51.795, 20.97926, 72.92726  (xmin, xmax, ymin,
ymax)
## coord. ref. : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84
+towgs84=0,0,0
## data source : in memory
## names       : Joint_Probability
## values      : 1.607938e-175, 0.07854836  (min, max)
```

Probability that at least one individual came from the location (union of probabilities)

```
unionP(asn)
```




```
## class       : RasterLayer
## dimensions  : 156, 339, 52884  (nrow, ncol, ncell)
## resolution  : 0.333, 0.333  (x, y)
## extent      : -164.682, -51.795, 20.97926, 72.92726  (xmin, xmax, ymin,
ymax)
## coord. ref. : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84
+towgs84=0,0,0
## data source : in memory
## names       : layer
## values      : 0.0006676678, 0.003915589  (min, max)
```