Trajectory Analysis for Samples Collected at the Yaqui Valley and Mexicali Valley sites March 2011 – June 2012

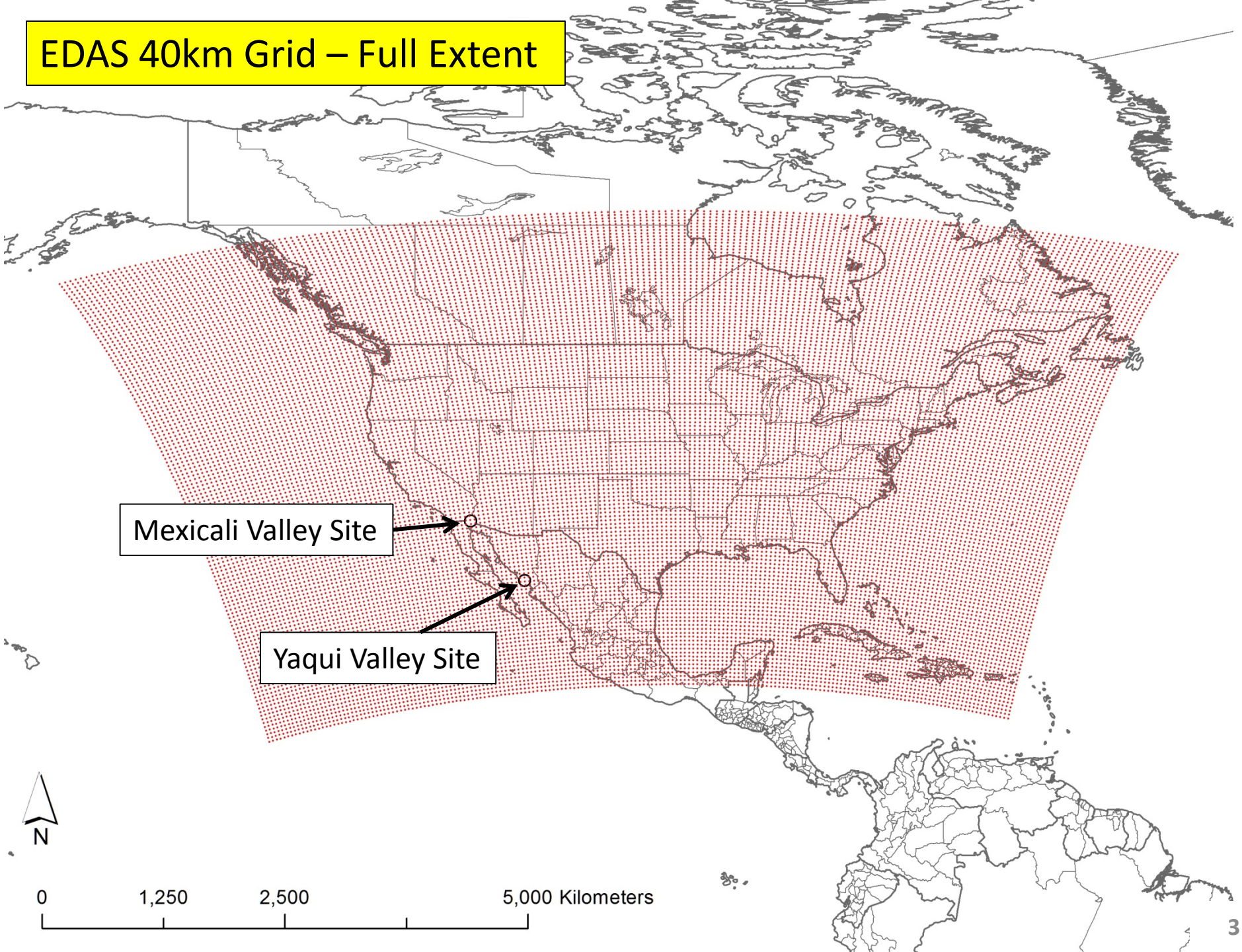
March 11, 2014

Mark Cohen
NOAA Air Resources Laboratory
College Park, Maryland, USA

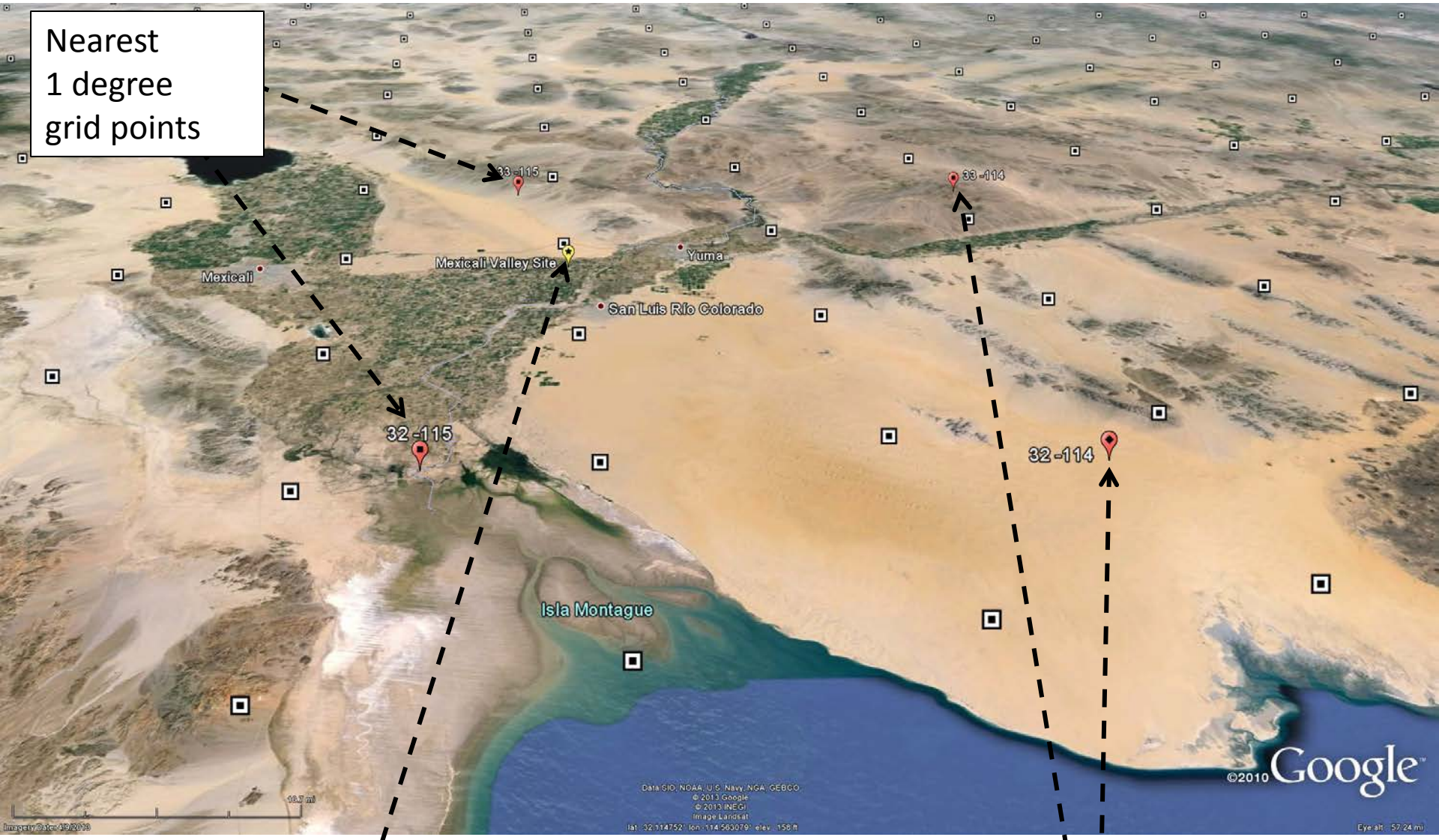
What “met data” to use to compute back-trajectories?

- First, tried to determine which meteorological data archive might be used
- The meteorological data are used to “drive” the HYSPLIT trajectory model
- We have some datasets converted to HYSPLIT format and archived here at NOAA ARL
- The following are datasets that are available for the region and time period of interest:
  - EDAS 40-km
  - GDAS (1 degree)
  - NCEP/NCAR Reanalysis (2.5 degree)
- Highest resolution grid in NOAA ARL archives grid covering the sampling region was the EDAS 40km dataset

EDAS 40km Grid – Full Extent



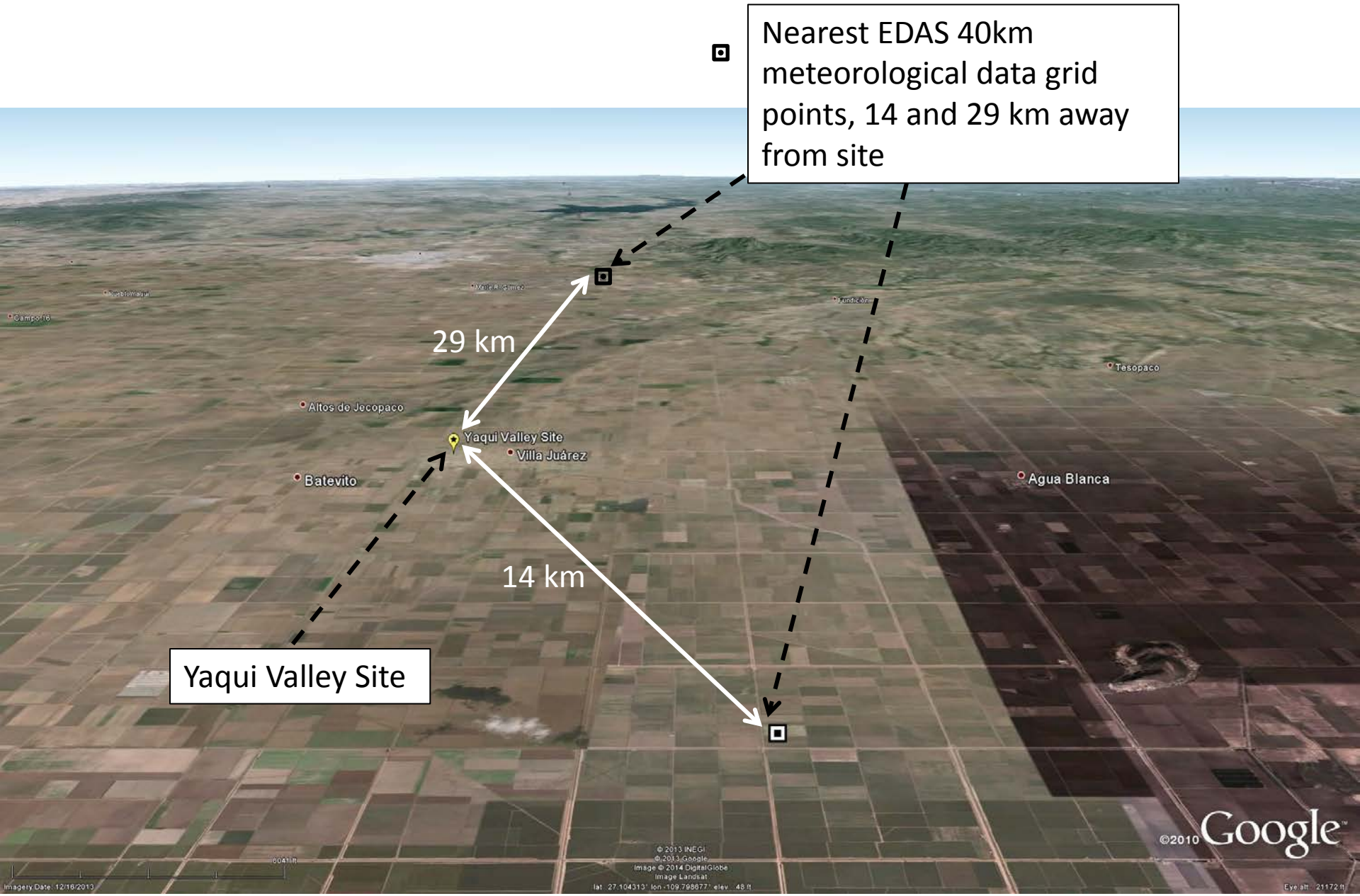
- EDAS 40km meteorological data grid points



Nearest EDAS 40km meteorological data grid point, 7 km away from site

Mexicali Valley Site

- EDAS 40km meteorological data grid points


Nearest EDAS 40km meteorological data grid points, 14 and 29 km away from site
29 km
14 km
Yaqui Valley Site
Altos de Jecopaco
Yaqui Valley Site
Villa Juárez
Batevito
Tesopaco
Agua Blanca
©2010 Google

BACK TRAJECTORY METHODOLOGY

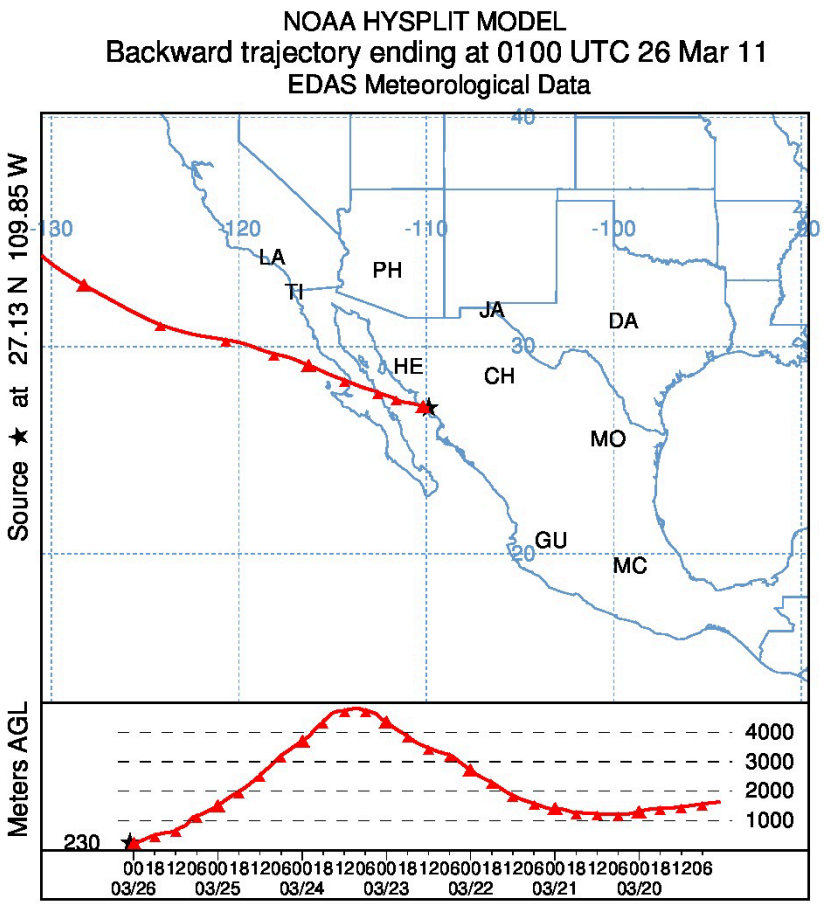
- NOAA-EDAS-40km meteorological data were used, after conversion to HYSPLIT format. Outside the domain of this dataset, the NCEP-NCAR Global Reanalysis (2.5 deg) meteorological data were used.

- Back-trajectories were simulated with the NOAA HYSPLIT model starting at a height of ½ of the Planetary Boundary Layer (PBL) height every 6 hours for the duration of a given sample: 12 AM, 6 AM, 12 PM, 6 PM (local time)

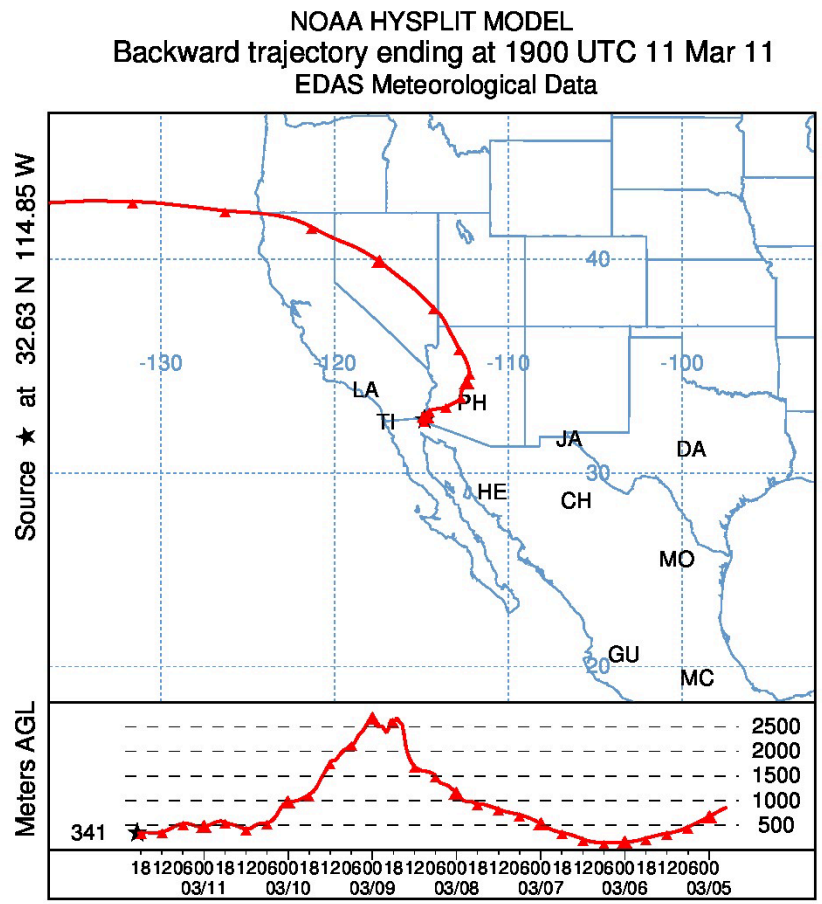
- Each trajectory was run for 168 hrs, and endpoints were output at a frequency of 30 minutes.

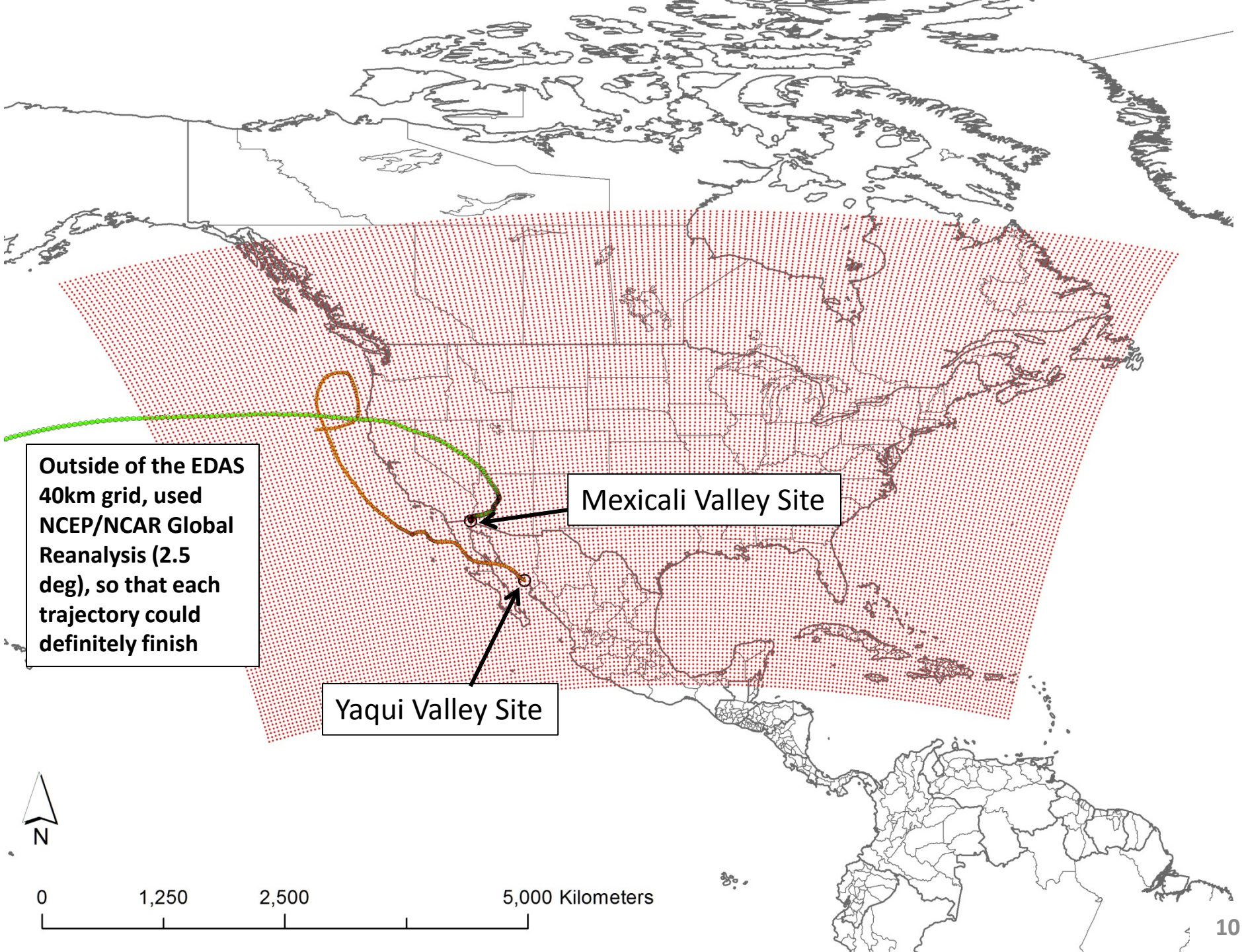
- So, 336 endpoints (lat, long, elevation) per back-trajectory

An example of a back-trajectory from the Yaqui Valley site, displayed using HYSPLIT graphics (TRAJPLOT)



An example of a back-trajectory from the Mexicali Valley site, displayed using HYSPLIT graphics (TRAJPLOT)




Outside of the EDAS 40km grid, used NCEP/NCAR Global Reanalysis (2.5 deg), so that each trajectory could definitely finish
Mexicali Valley Site
Yaqui Valley Site
N
0
1,250
2,500
5,000 Kilometers

Sample	Time Period	Number of Days	Number of Trajectories (4 per day)	Number of Endpoints (336 per each 168 hr trajectory)
Yaqui Valley T1	25 Mar 2011 – 1 July 2011	98	392	131,712
Yaqui Valley T2	1 July 2011 – 7 Oct 2011	98	392	131,712
Yaqui Valley T3	7 Oct 2011 – 12 Feb 2012	128	512	172,032
Yaqui Valley T4	12 Feb 2012 – 2 June 2012	131	444	149,184
Yaqui Valley Total	***25 Mar 2011 – 2 June 2012***	***435***	***1,740***	***584,640***
Mexicali Valley T1	11 Mar 2011 – 11 June 2011	92	368	123,648
Mexicali Valley T2	11 June 2011 – 11 Oct 2011	122	488	163,968
Mexicali Valley T3	11 Oct 2011 – 12 Feb 2012	124	496	166,656
Mexicali Valley T4	12 Feb 2012 – 12 May 2012	90	360	120,960
Mexicali Valley Total	***11 Mar 2011 – 12 May 2012***	***428***	***1,712***	***575,232***

TRAJECTORY FREQUENCY ANALYSIS

- The set of trajectories corresponding to any given sample (or group of samples) were then subjected to a "grid-frequency" analysis, i.e., using the program "TRAJFREQ" (part of the HYSPLIT modeling suite)

- In this analysis, the number of 30-minute trajectory endpoints in each specified grid cell were added up.

- For the trajectory frequency results presented here, a 0.25 x 0.25 degree grid (about 25km x 25km) was used, to examine regional influences.

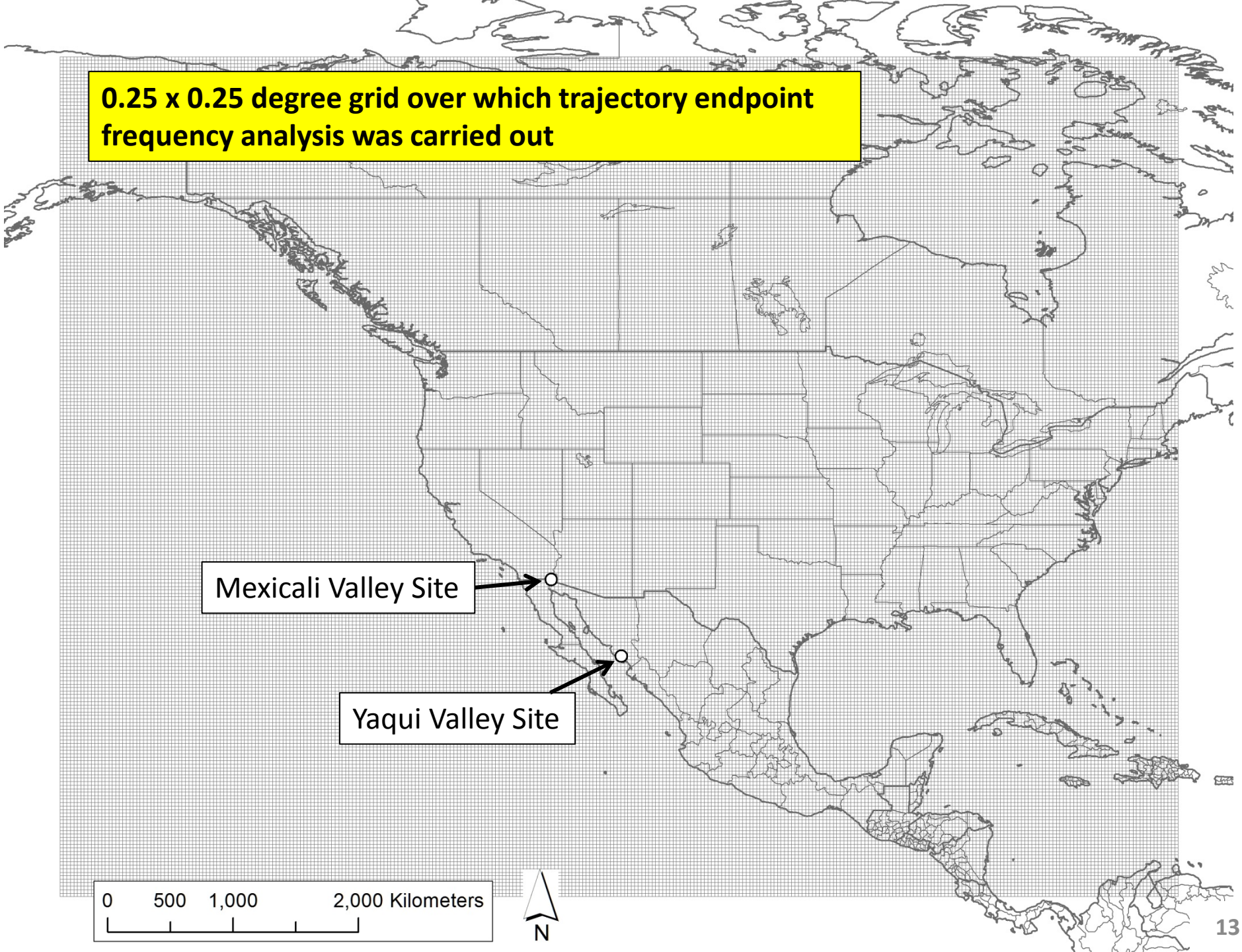
- The gridded trajectory frequency results mapped represent the following:

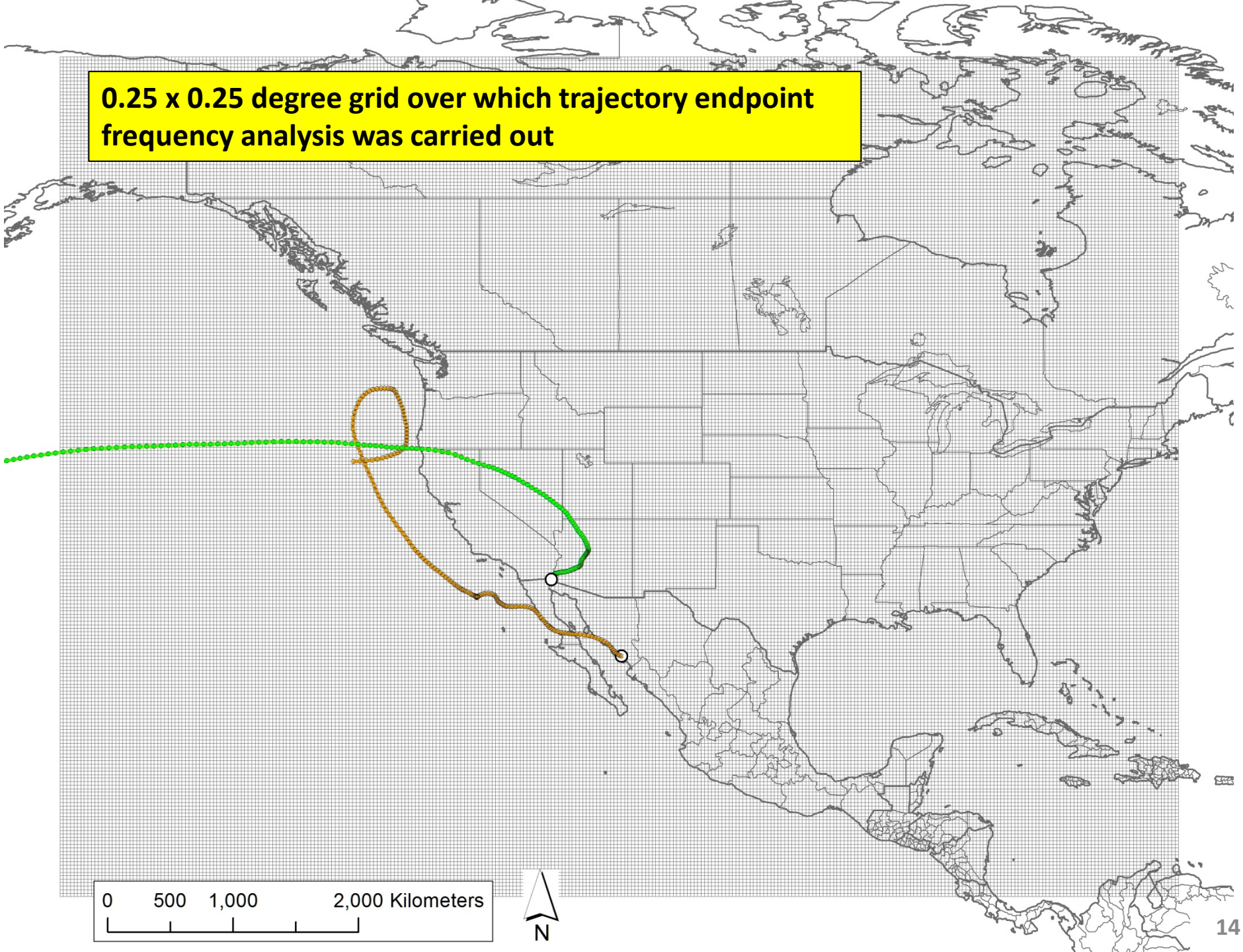
  FREQ = 100 * (number of endpoints per grid square / number of trajectories)

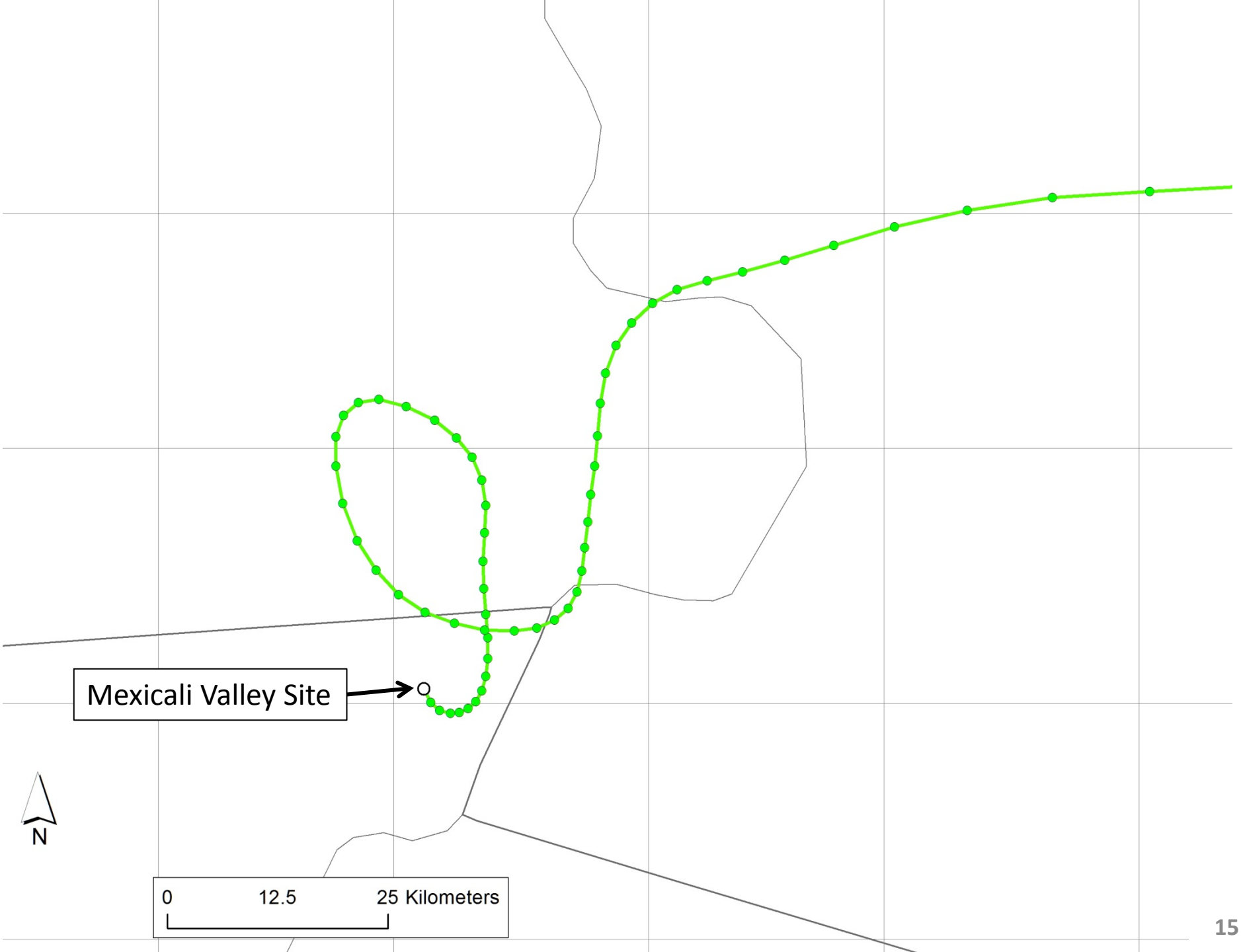
- So, for example, a grid cell value of "100" means that on average, one 30-minute endpoint from each trajectory fell in the grid cell. A grid cell value of "50" means that on average, one 30-minute endpoint from only 50% of the trajectories fell in the grid cell.

- Note that absolute values of these results are strongly influenced by the grid size used (in this case, 0.25 x 0.25 deg) and the endpoint output frequency (in this case, 30 minutes)

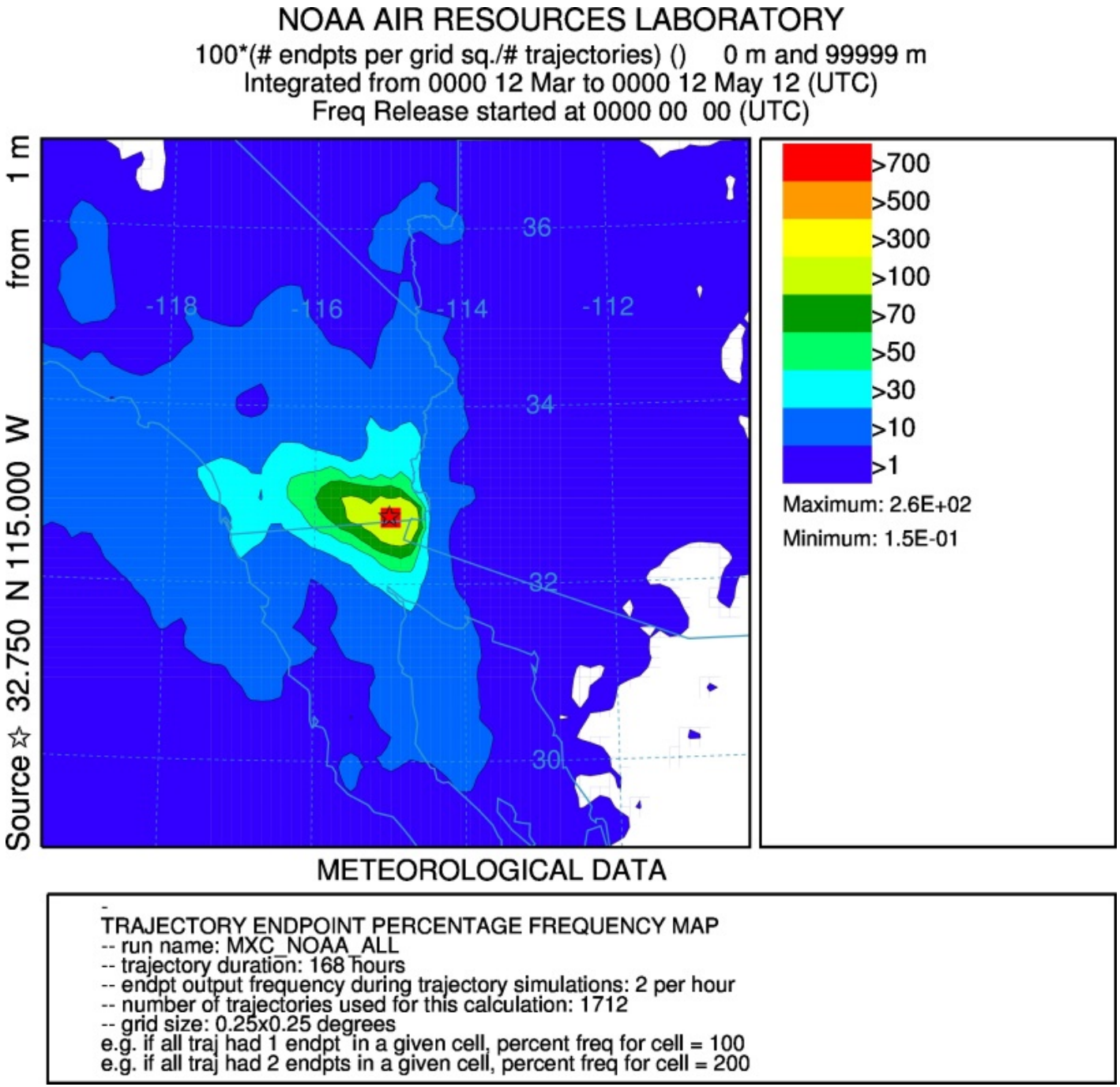
- However, the spatial patterns in the results are less strongly influenced by these factors.


0.25 x 0.25 degree grid over which trajectory endpoint frequency analysis was carried out
Mexicali Valley Site
Yaqui Valley Site
0 500 1,000 2,000 Kilometers
N

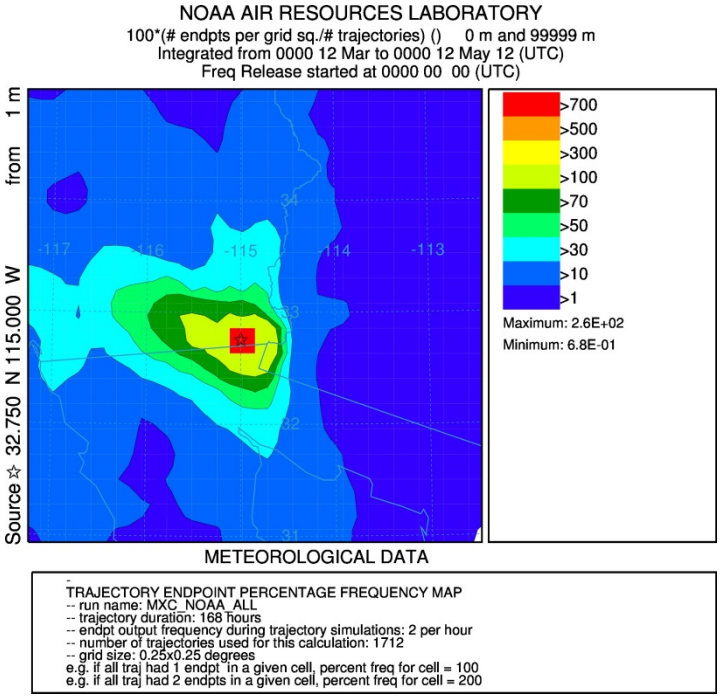

0.25 x 0.25 degree grid over which trajectory endpoint frequency analysis was carried out
0 500 1,000 2,000 Kilometers
N

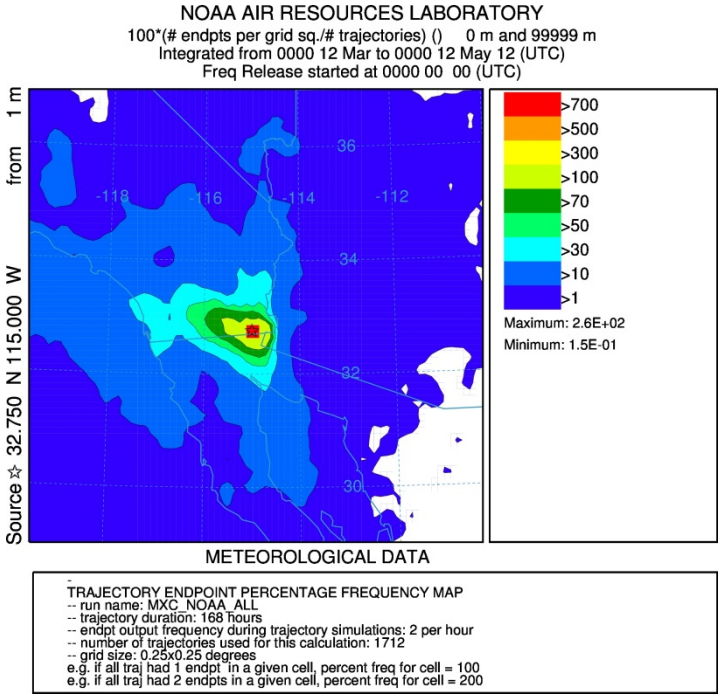

Mexicali Valley Site
N
0
12.5
25 Kilometers

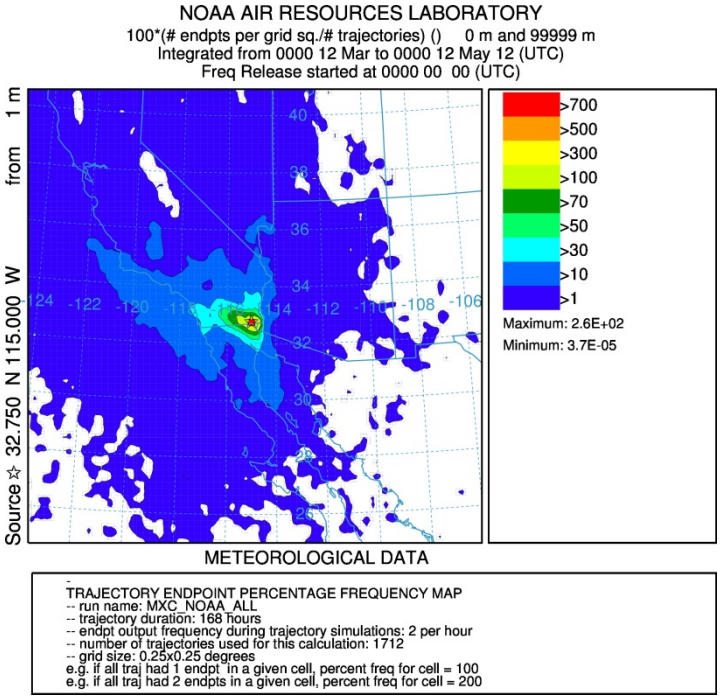
Here's some output, using the HYSPLIT modeling suite graphics capability…

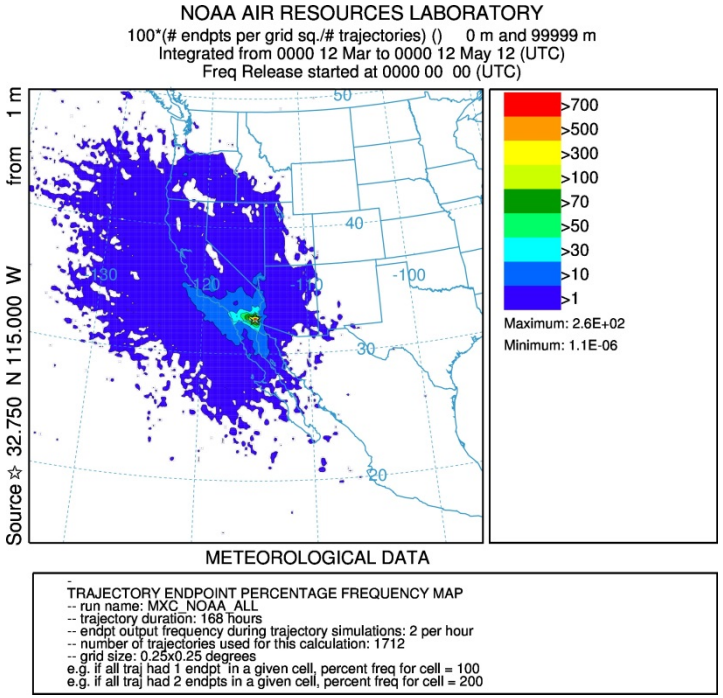


Here are the results for all of the trajectories run for the Mexicali Valley site, shown using four different scales

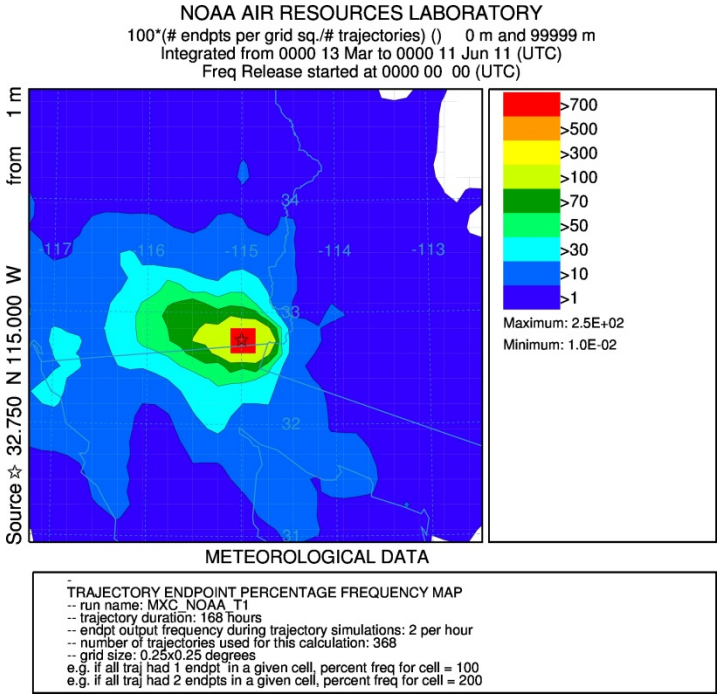


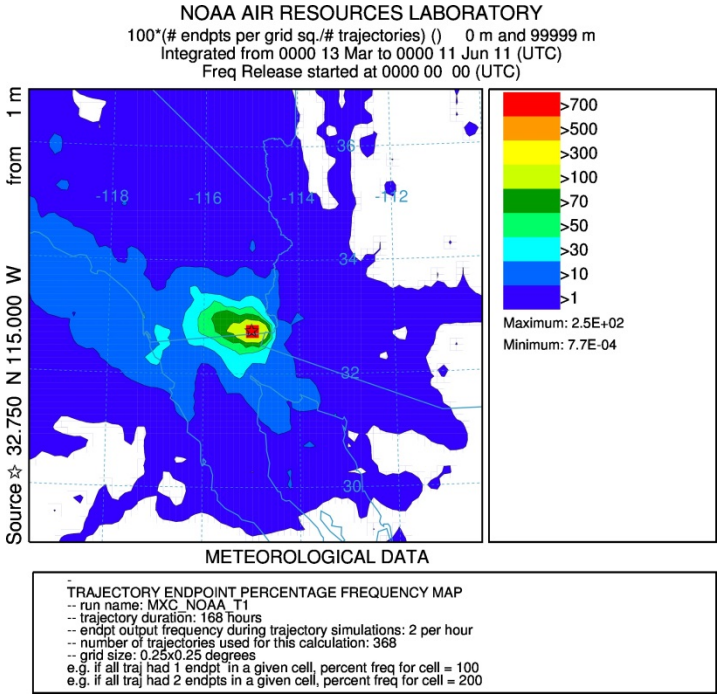


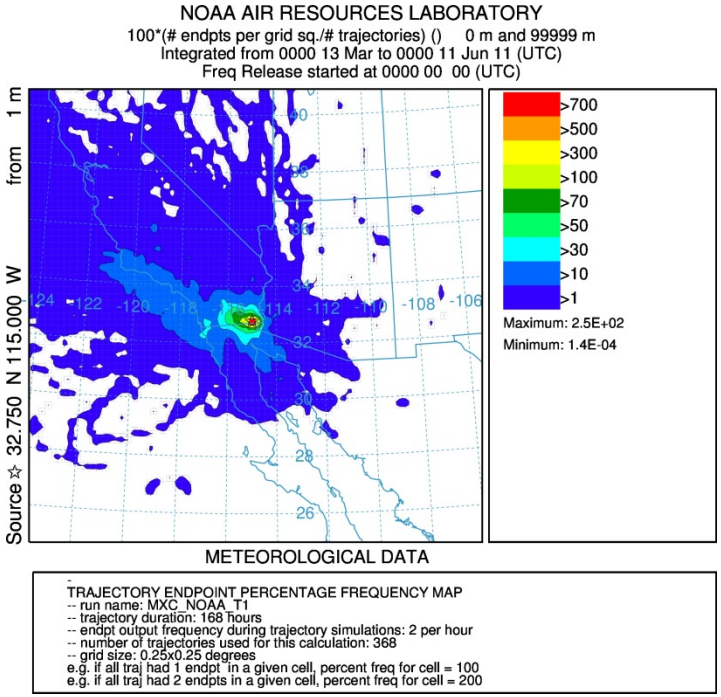


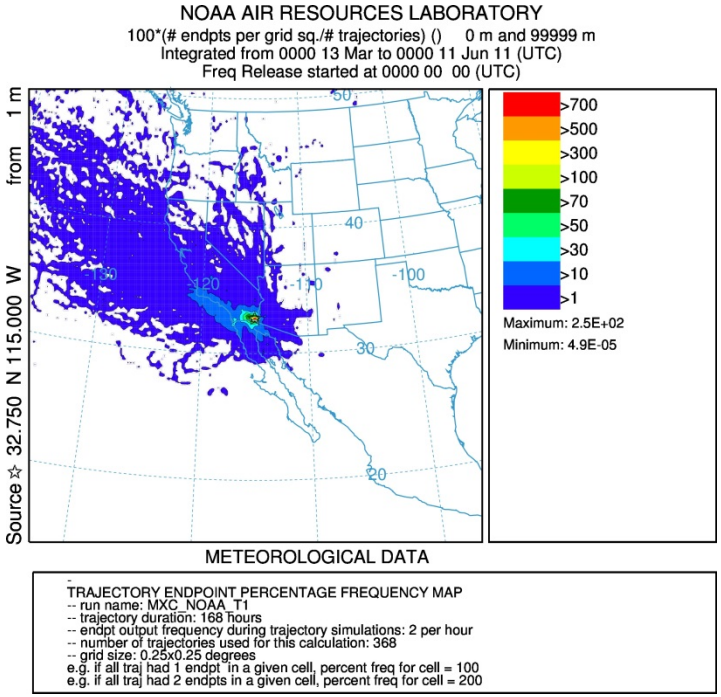


Here are the results for the trajectories corresponding to the Mexicali Valley site sample T1, shown using four different scales

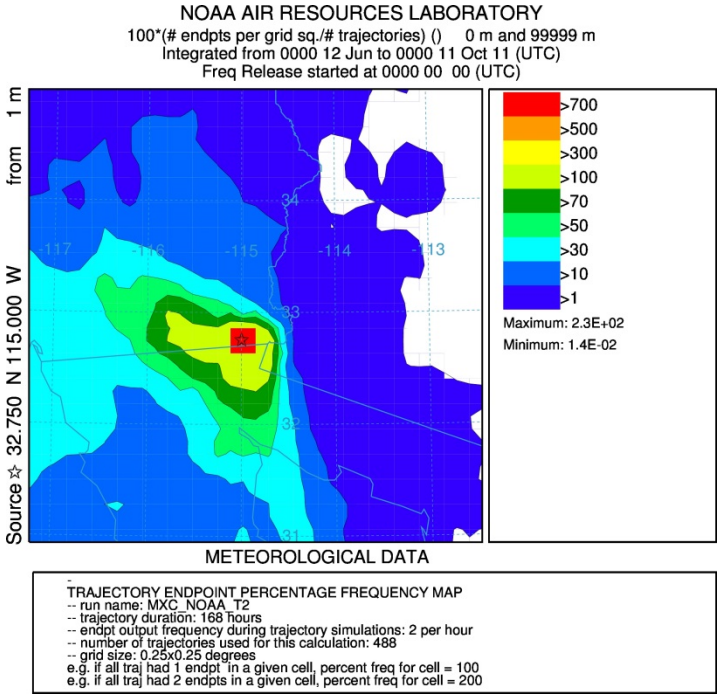


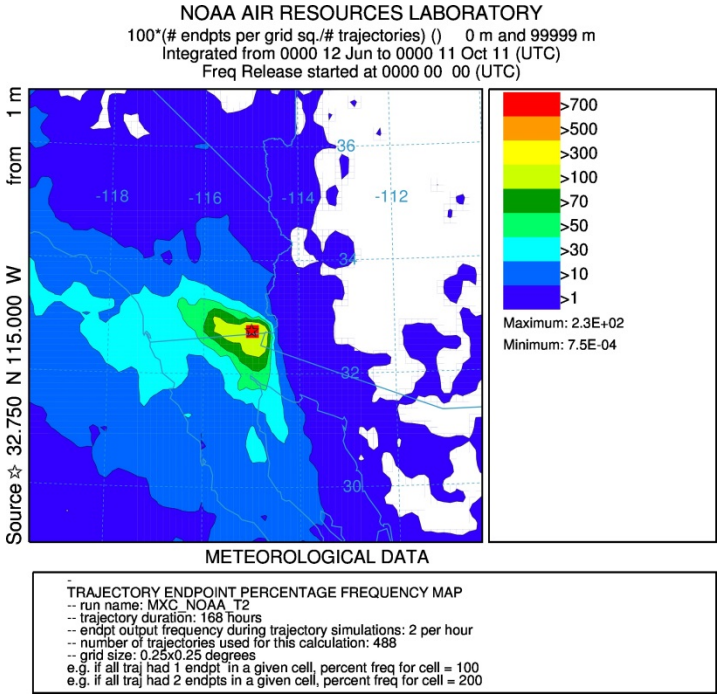


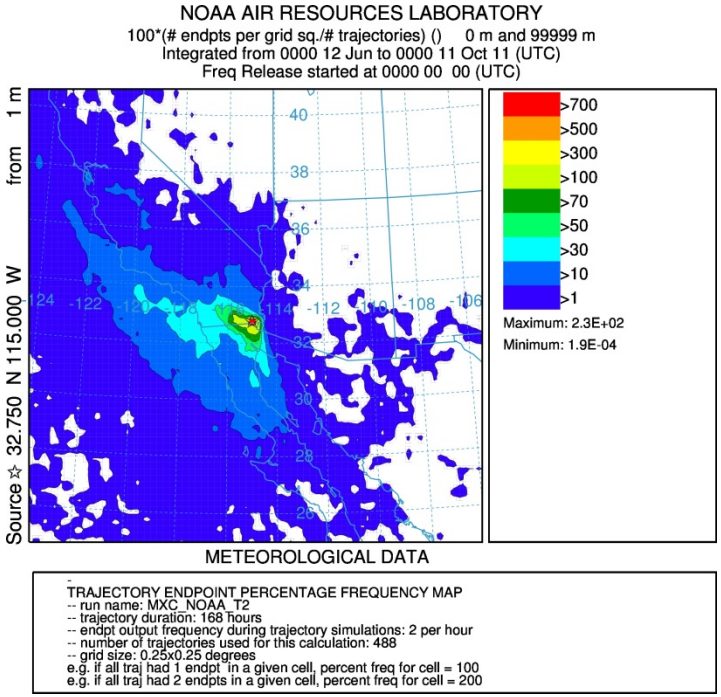


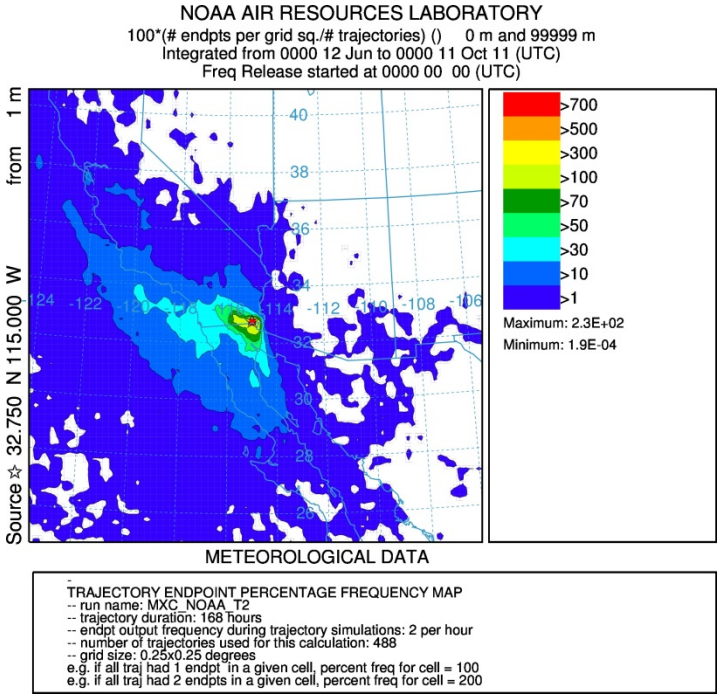


Here are the results for the trajectories corresponding to the Mexicali Valley site sample T2, shown using four different scales

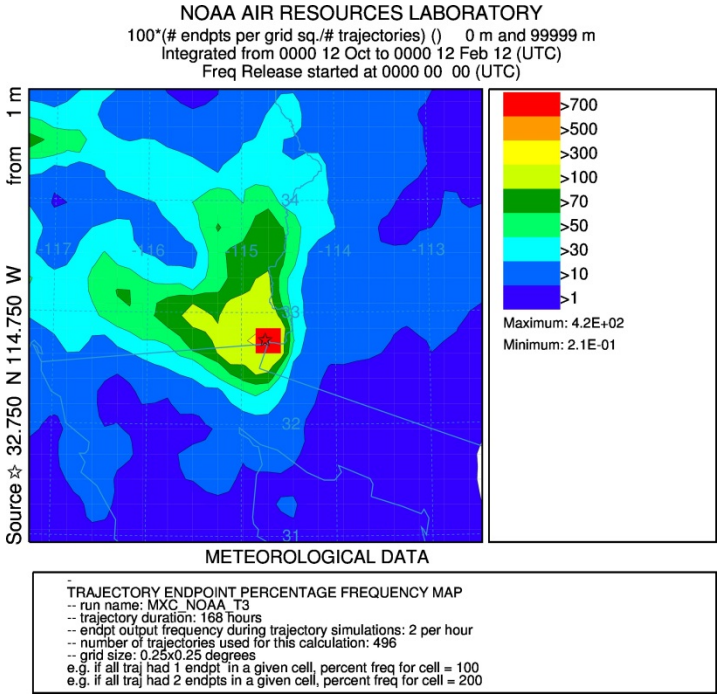


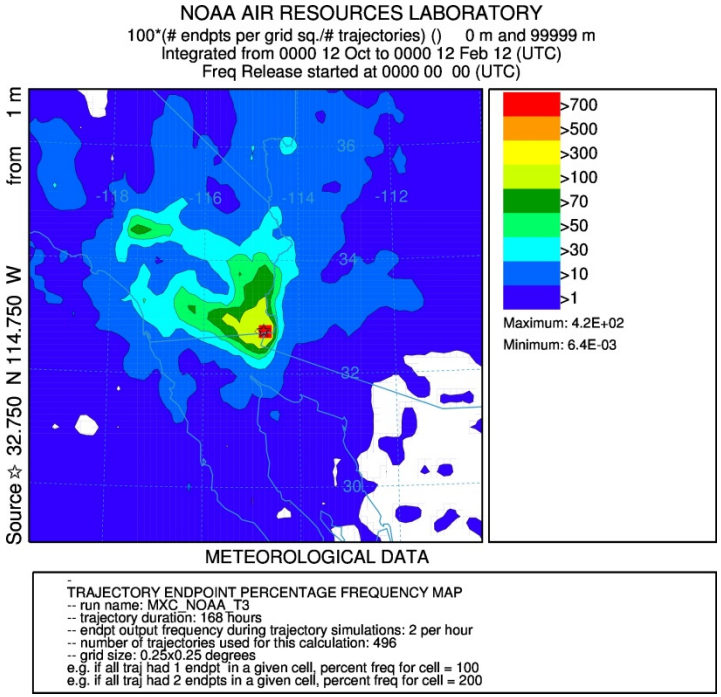


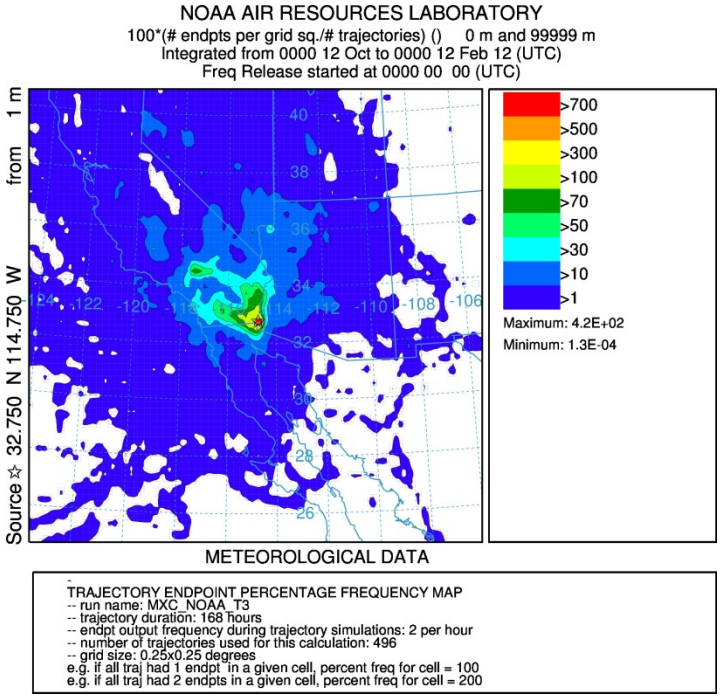


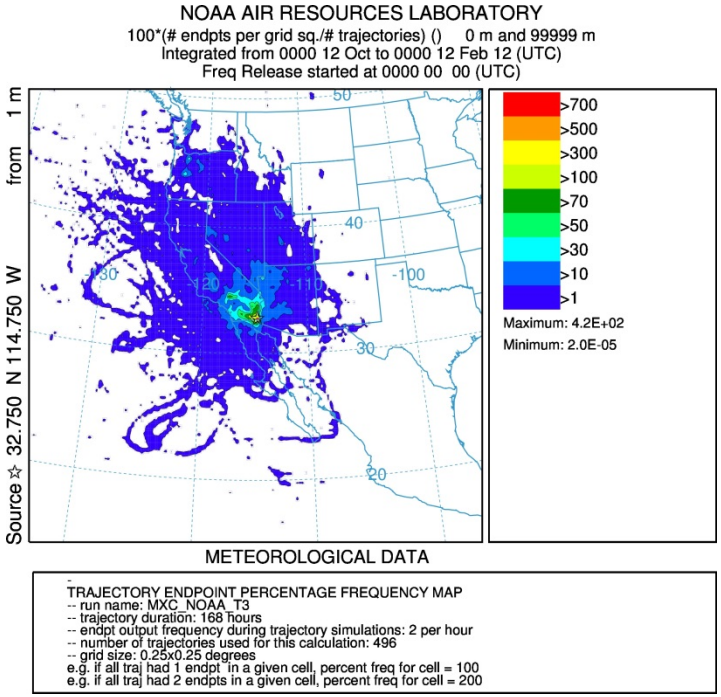


Here are the results for the trajectories corresponding to the Mexicali Valley site sample T3, shown using four different scales

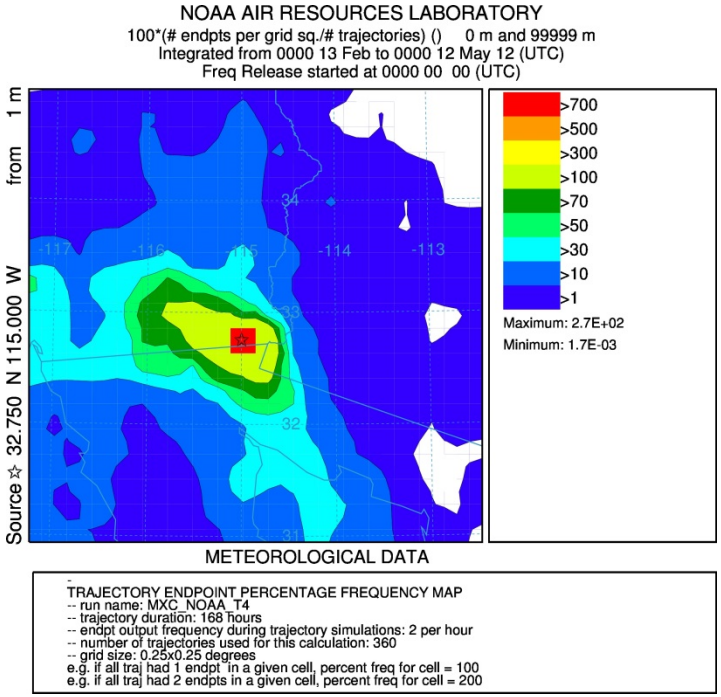


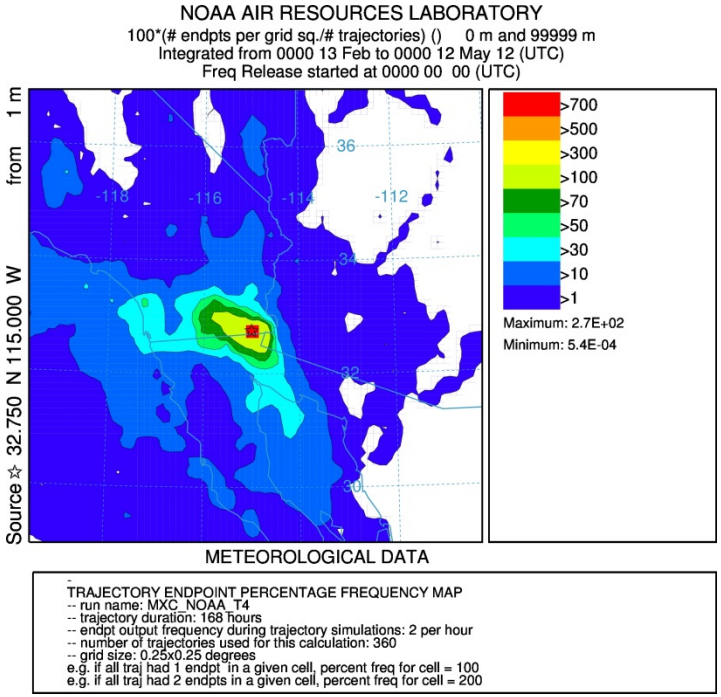


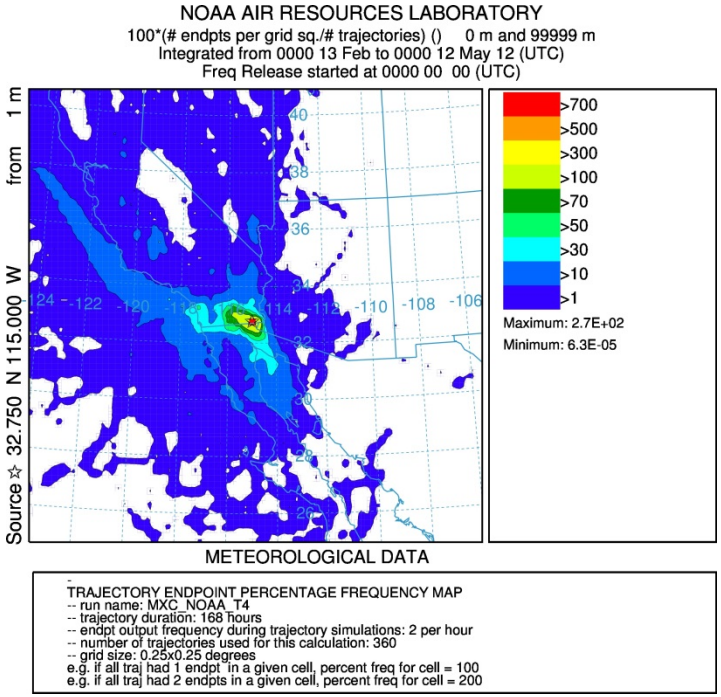


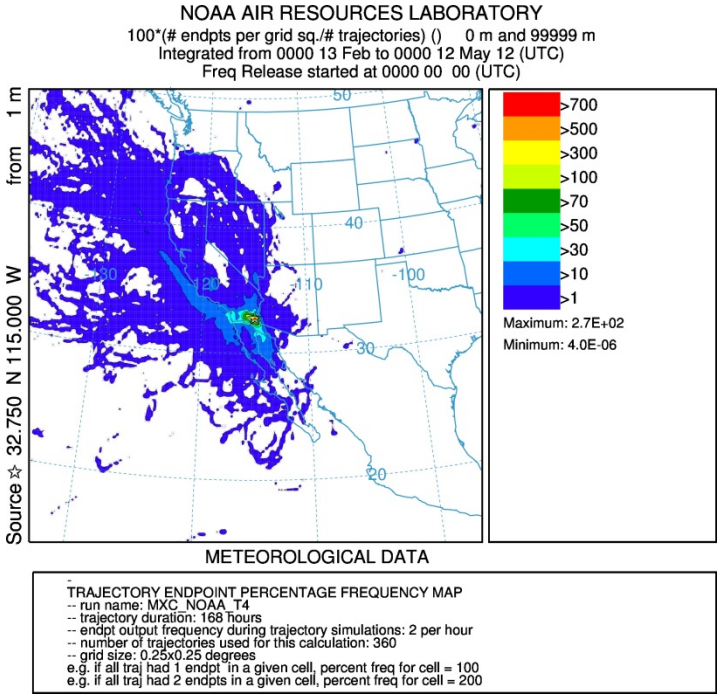


Here are the results for the trajectories corresponding to the Mexicali Valley site sample T4, shown using four different scales



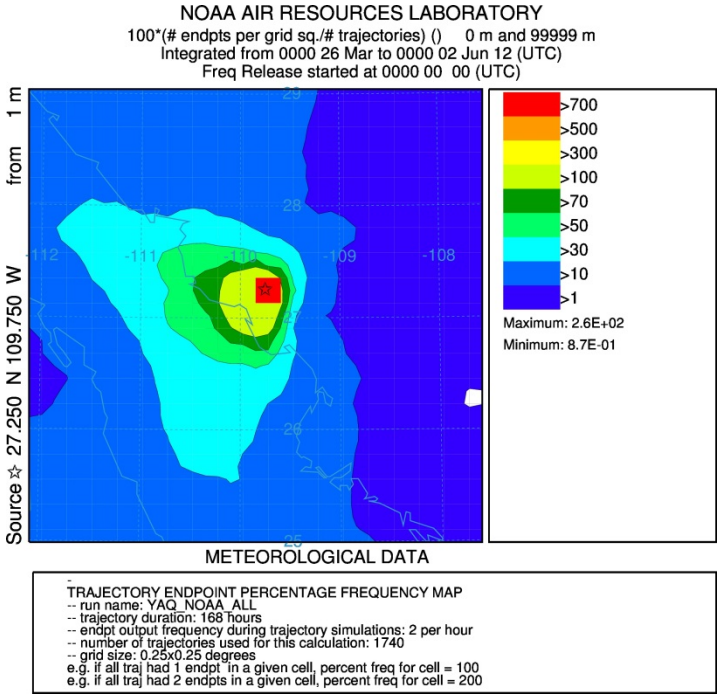


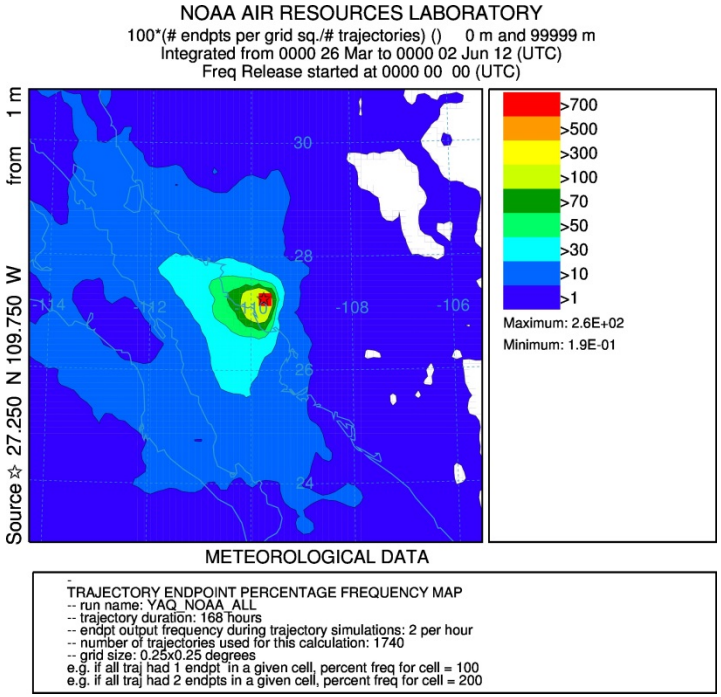


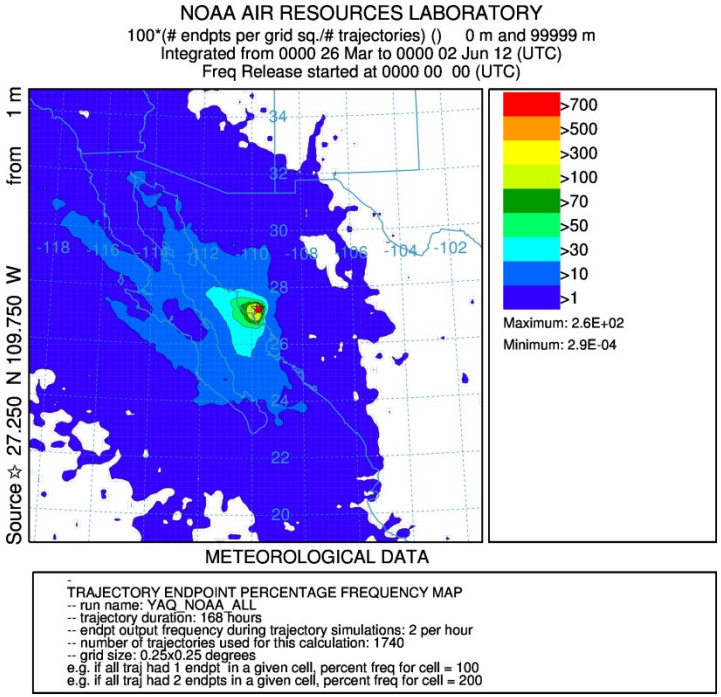


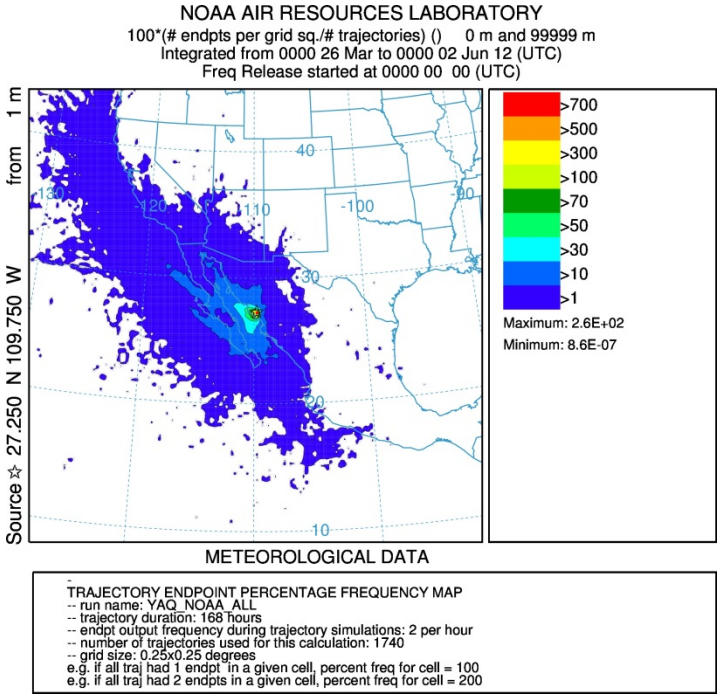
And now the same graphics for the Yaqui Valley site…

Here are the results for all of the trajectories run for the Yaqui Valley site, shown using four different scales

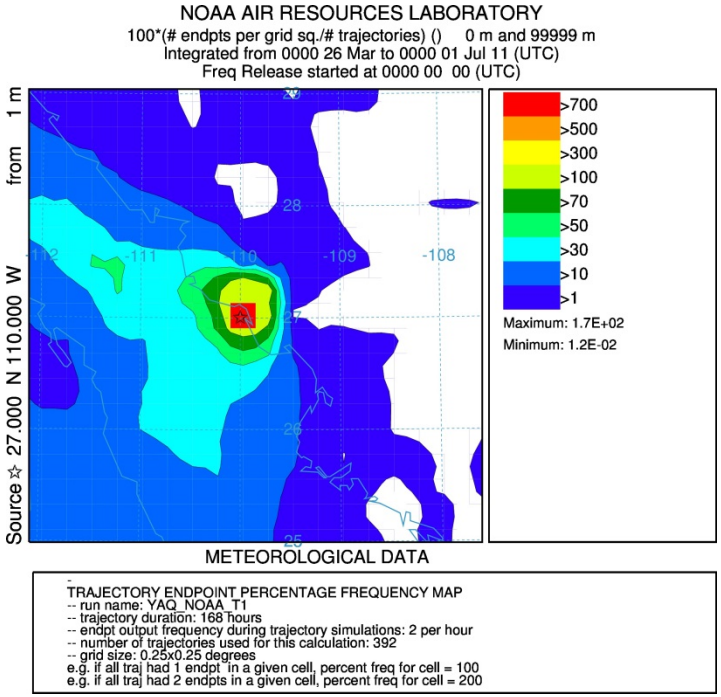


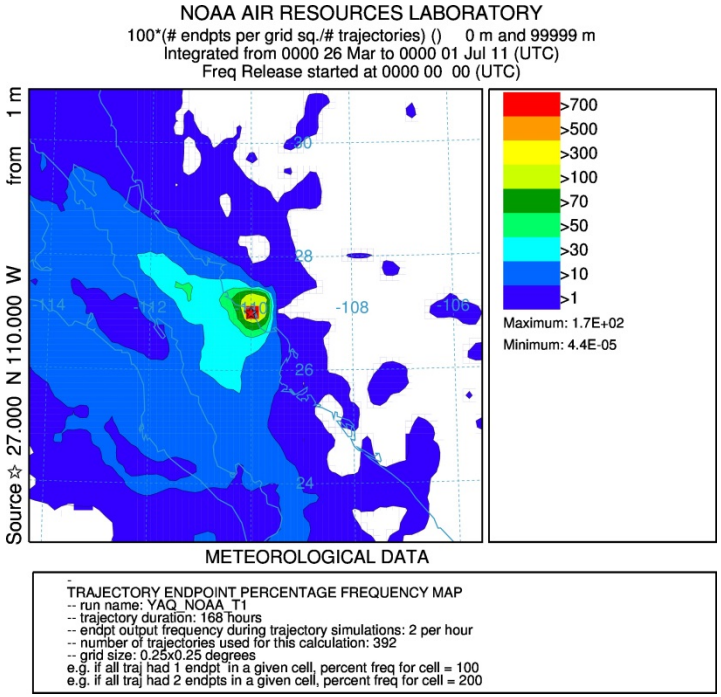


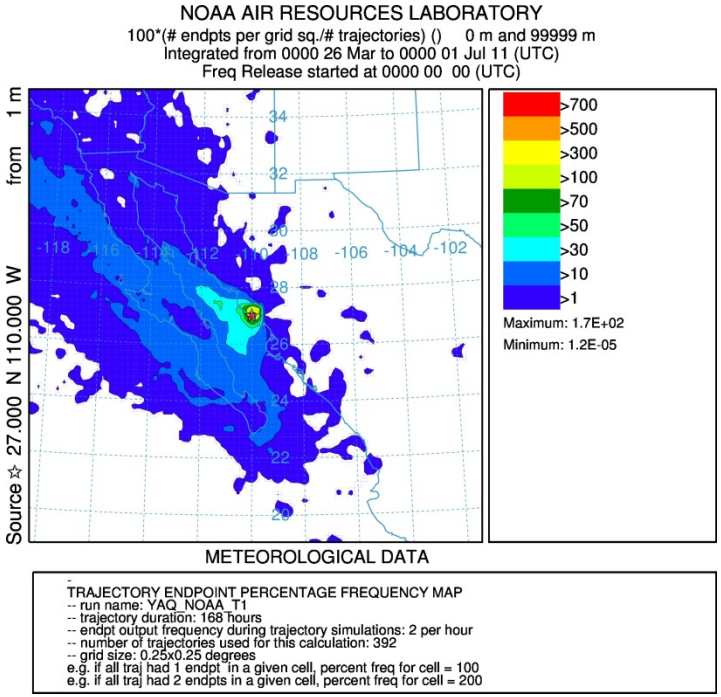


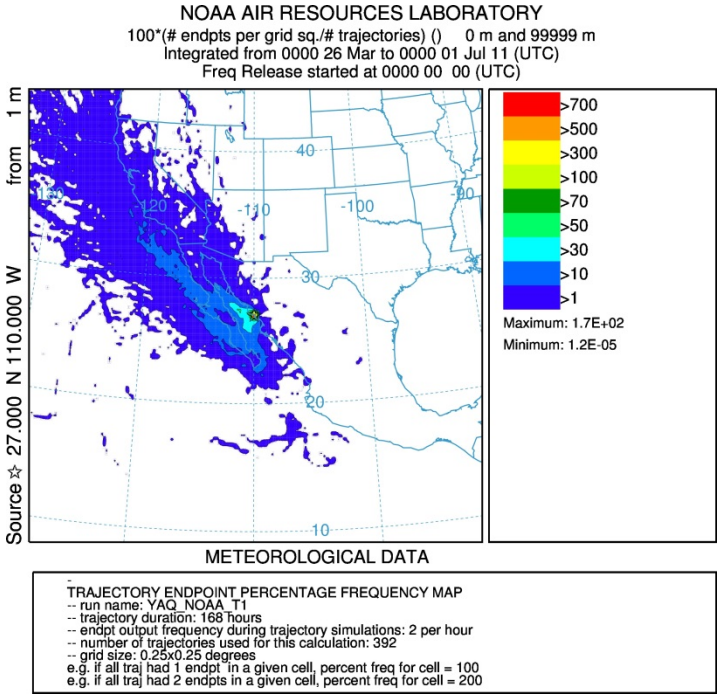


Here are the results for the trajectories corresponding to the Yaqui Valley site sample T1, shown using four different scales

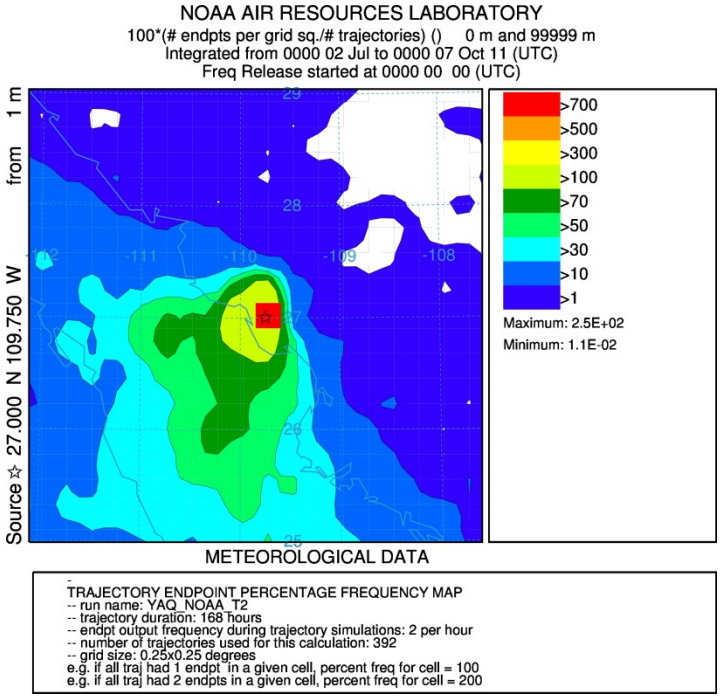


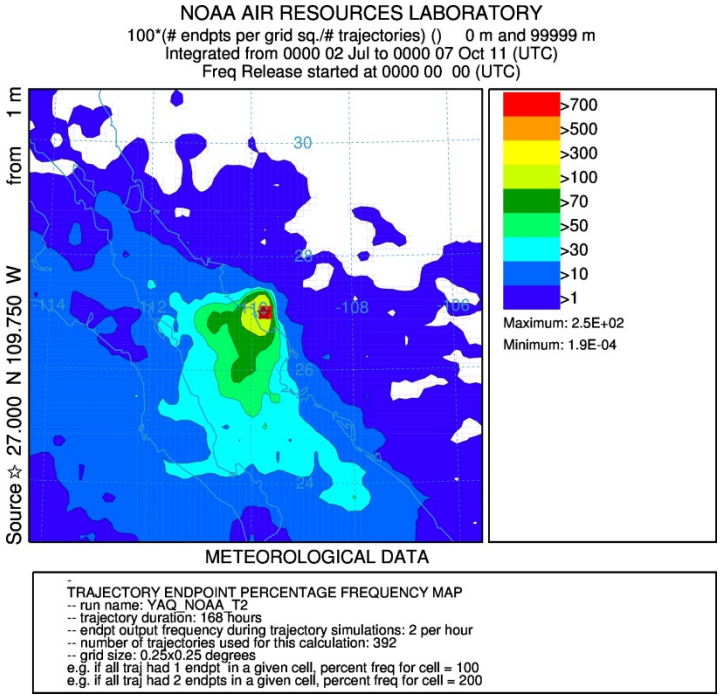


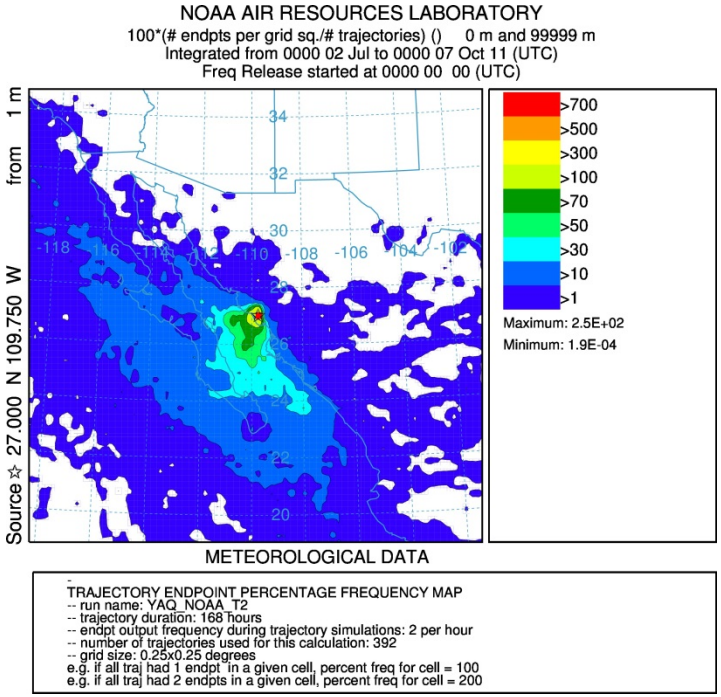


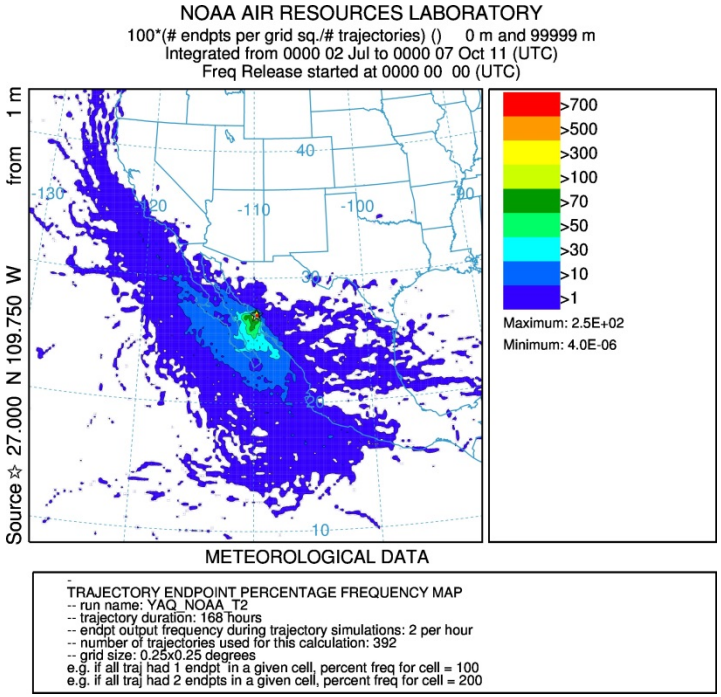


Here are the results for the trajectories corresponding to the Yaqui Valley site sample T2, shown using four different scales

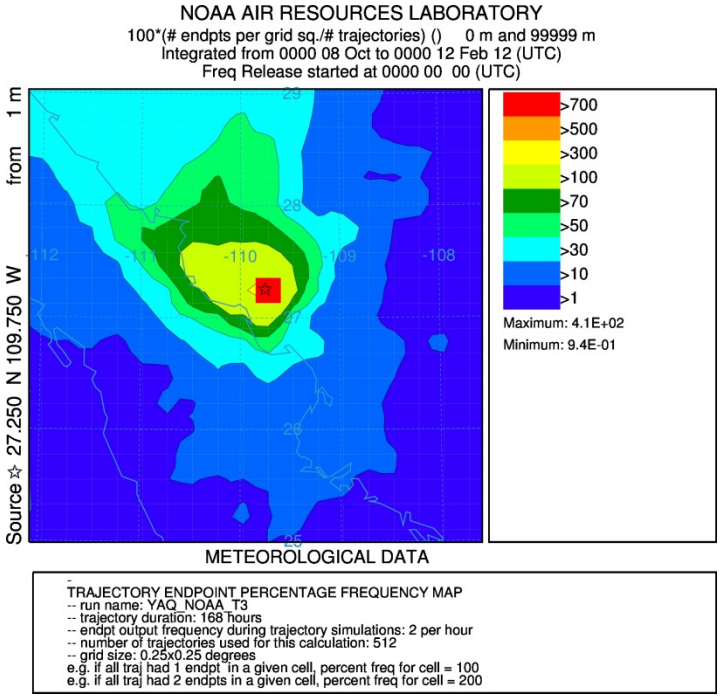


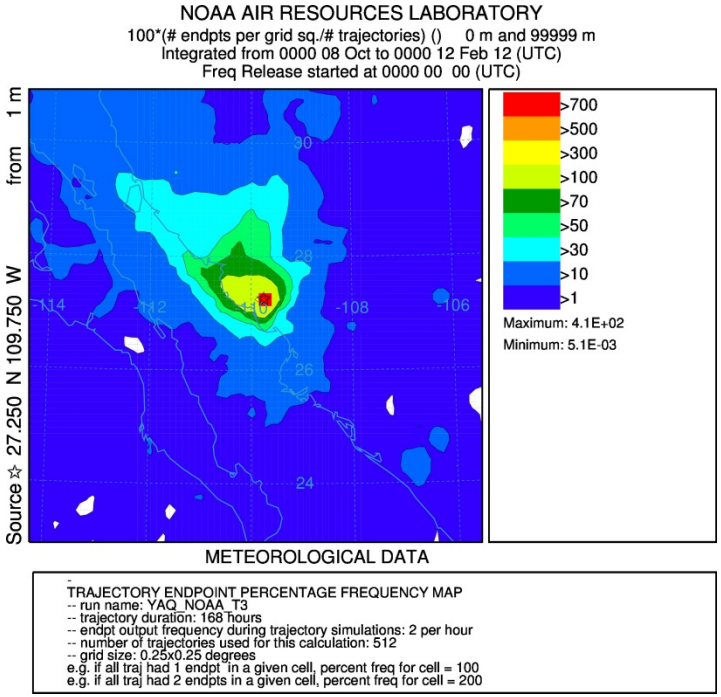


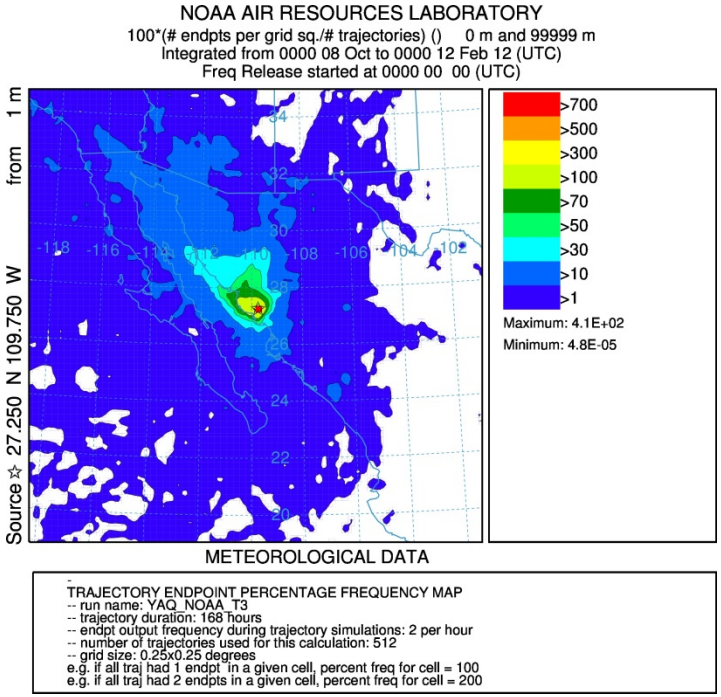


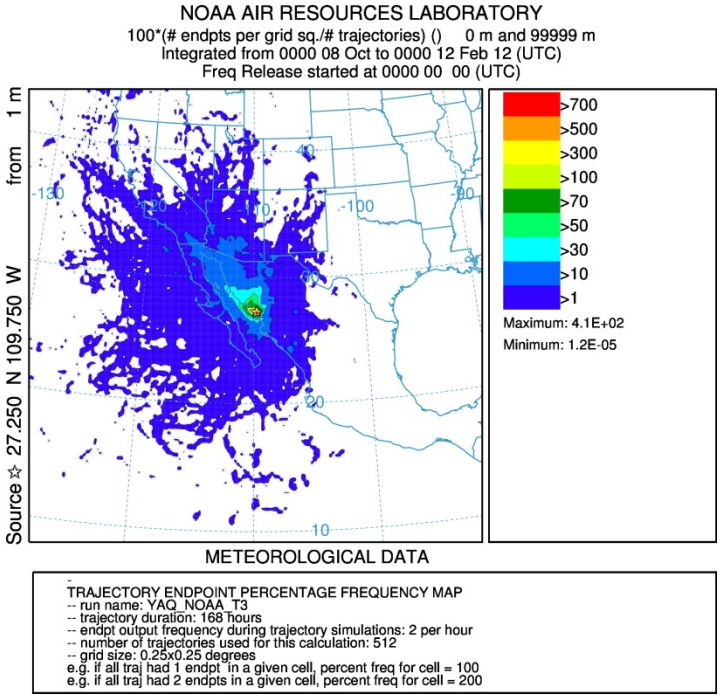


Here are the results for the trajectories corresponding to the Yaqui Valley site sample T3, shown using four different scales

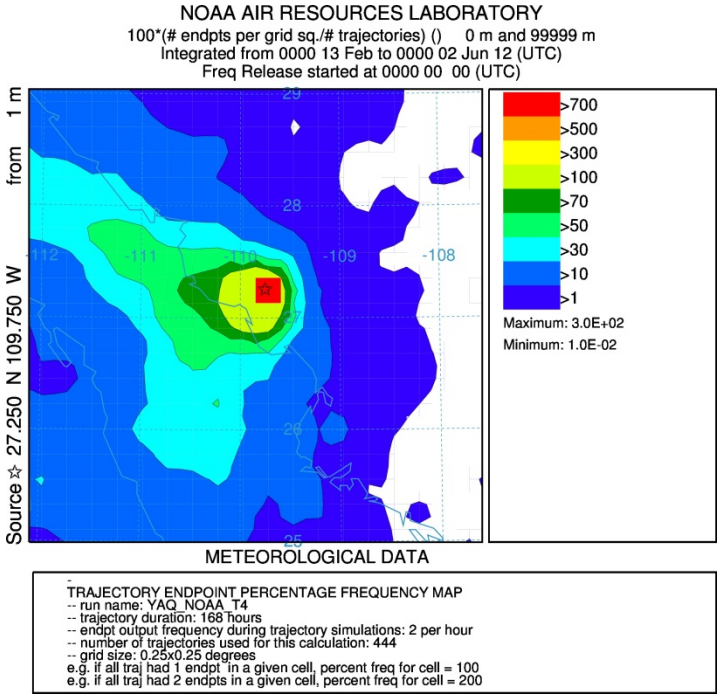


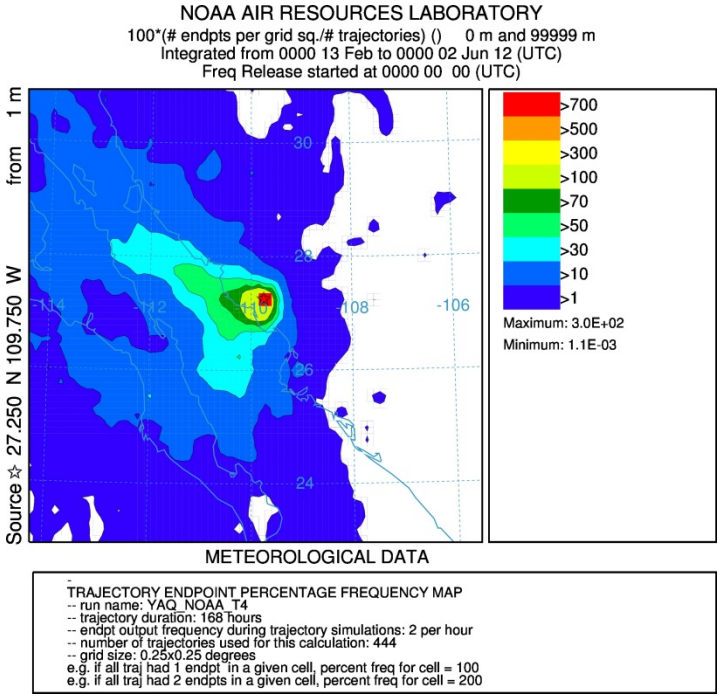


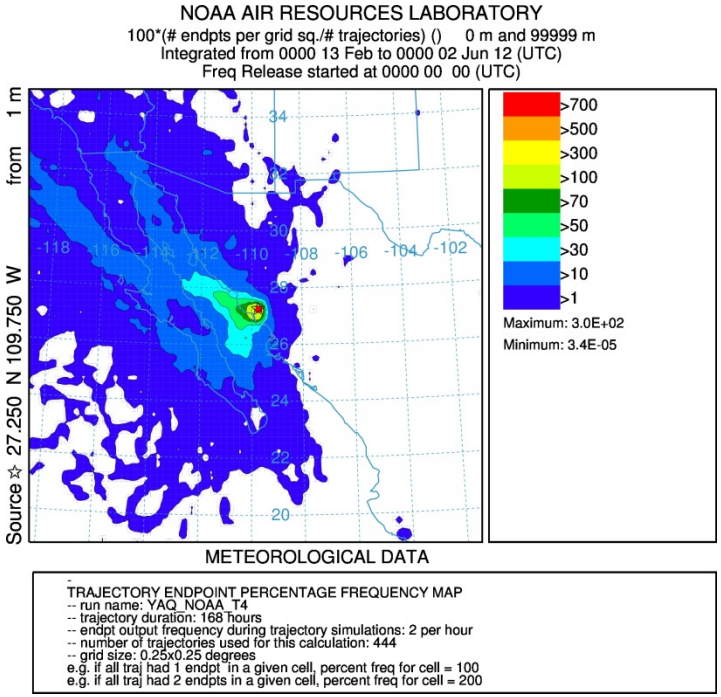


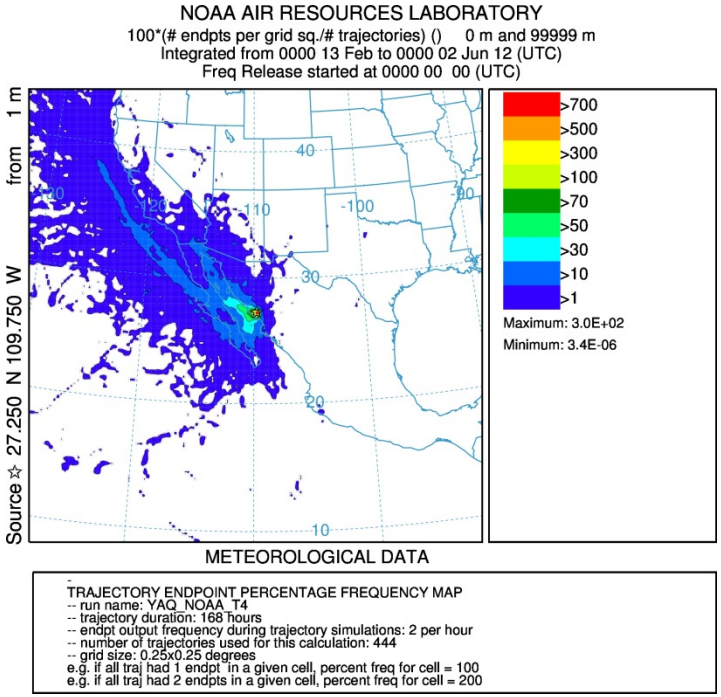


Here are the results for the trajectories corresponding to the Yaqui Valley site sample T4, shown using four different scales



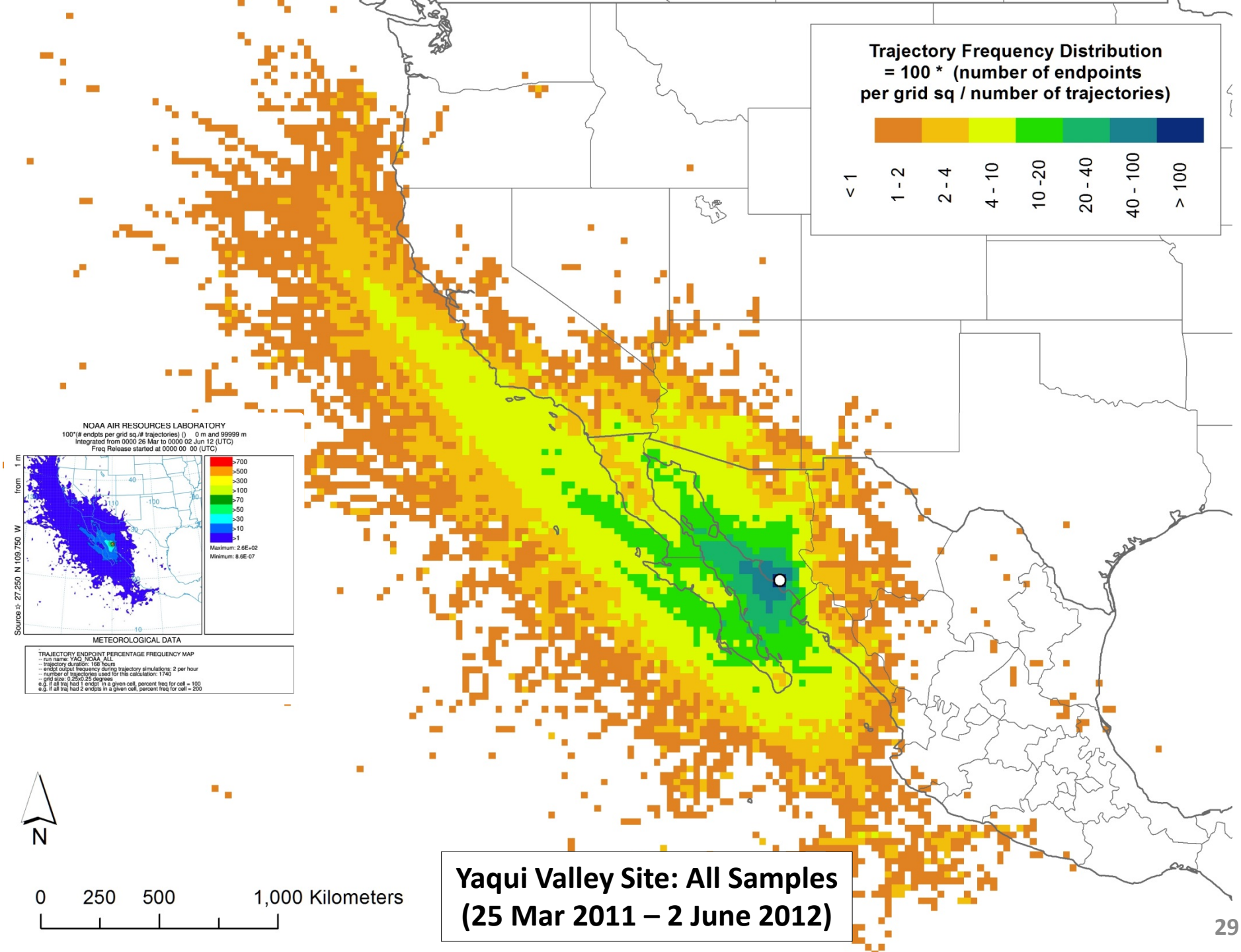




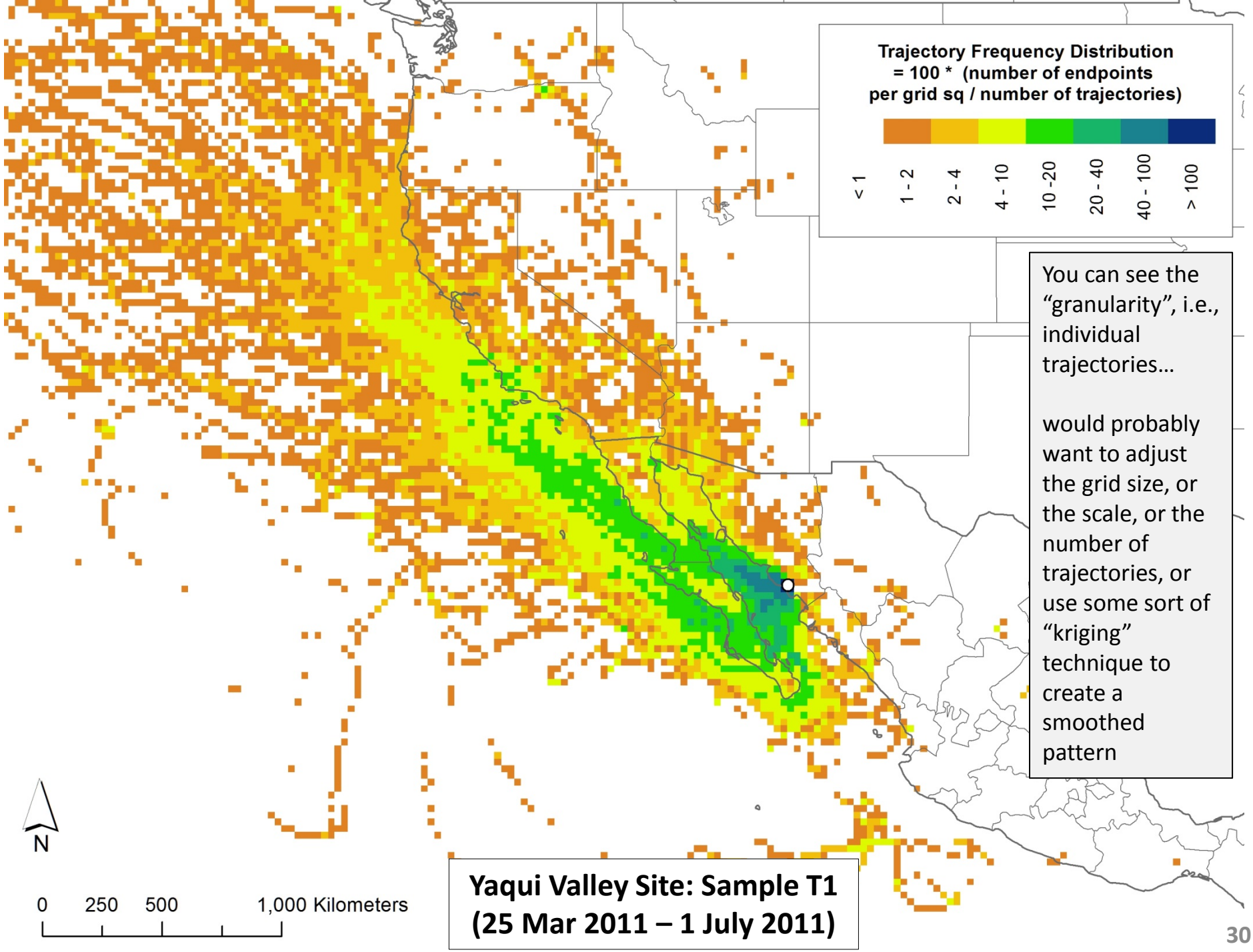


Can import the data into a GIS program…

Here are some examples using ESRI ArcMap



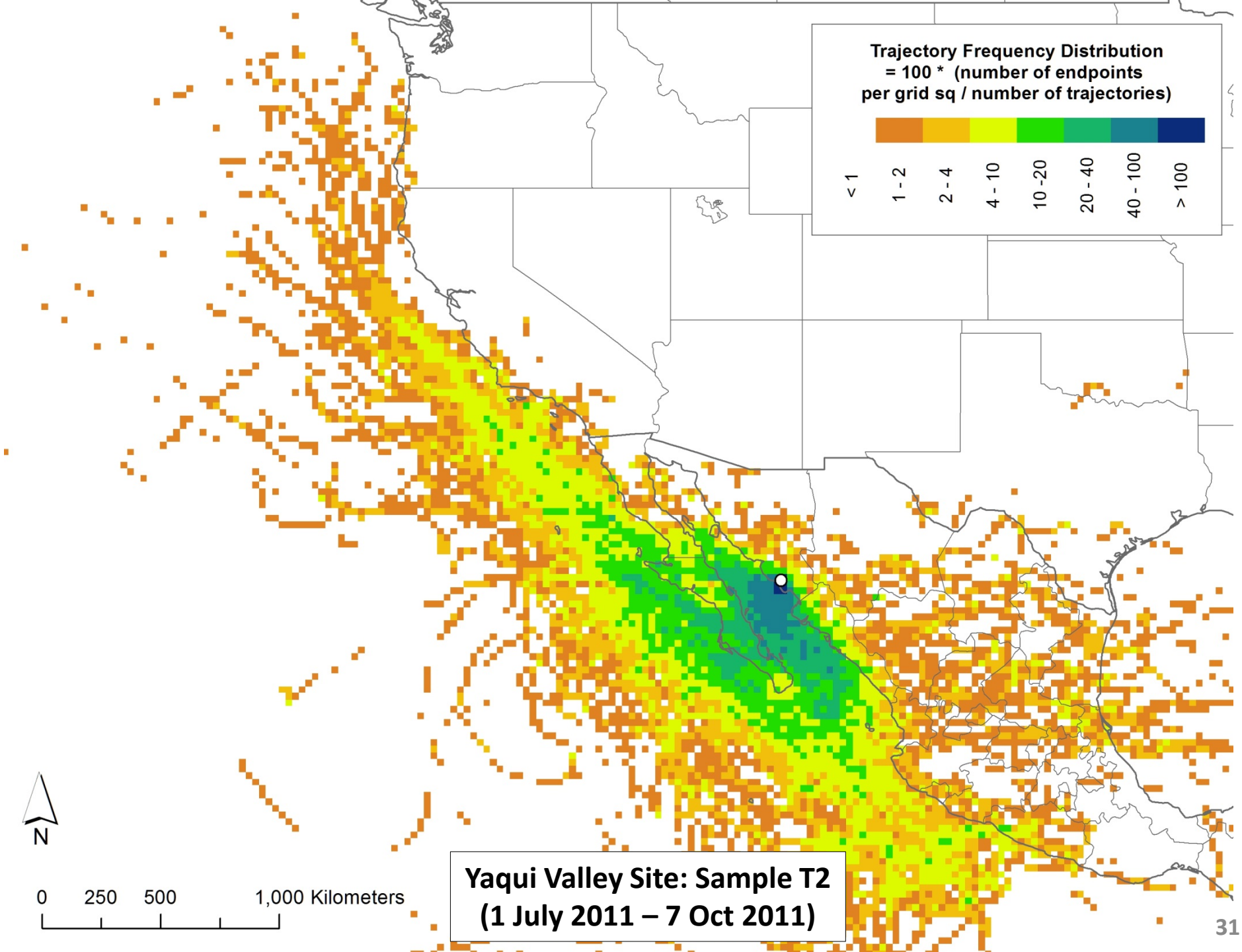
Yaqui Valley Site: All Samples (25 Mar 2011 – 2 June 2012)



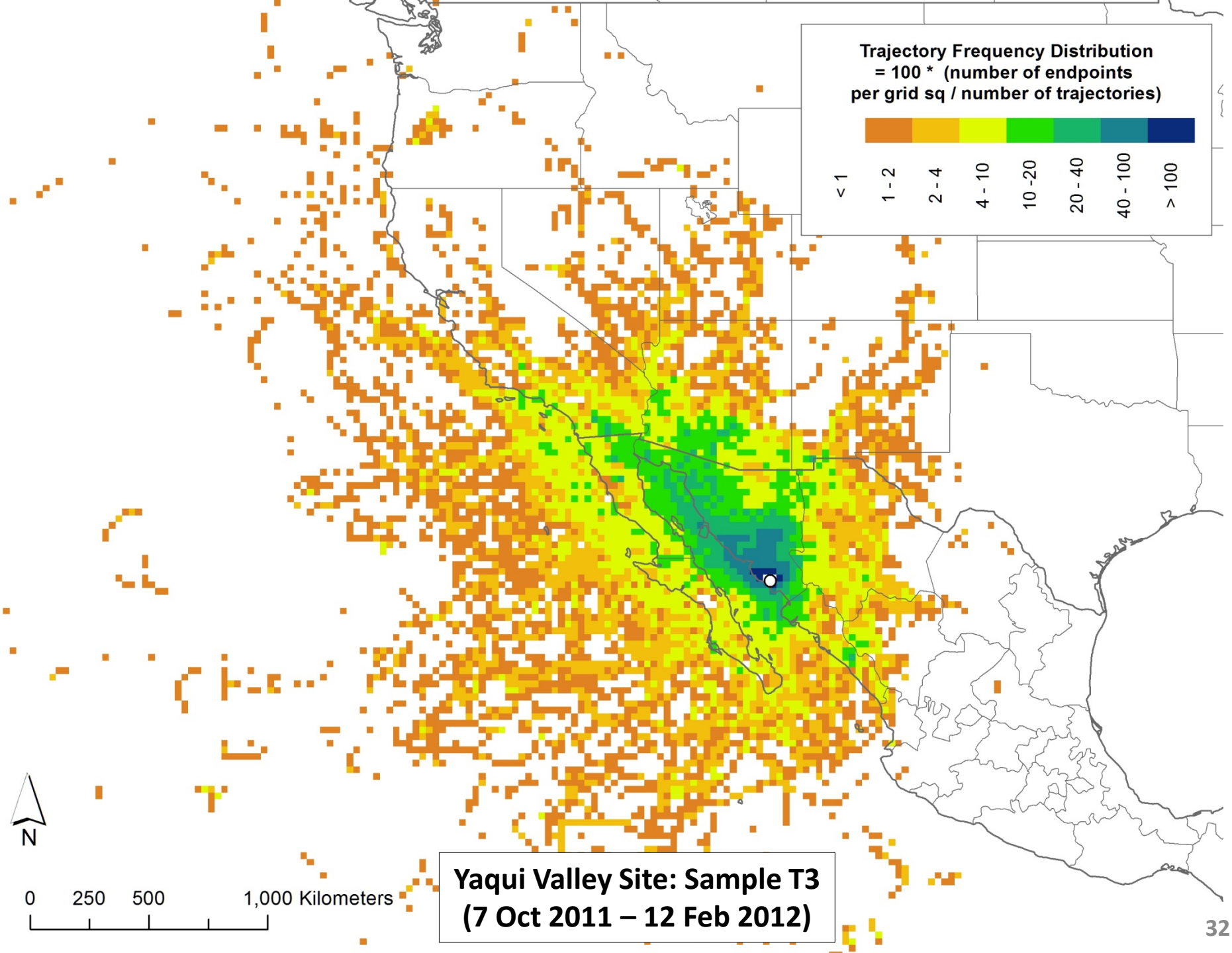
You can see the "granularity", i.e., individual trajectories…

would probably want to adjust the grid size, or the scale, or the number of trajectories, or use some sort of "kriging" technique to create a smoothed pattern

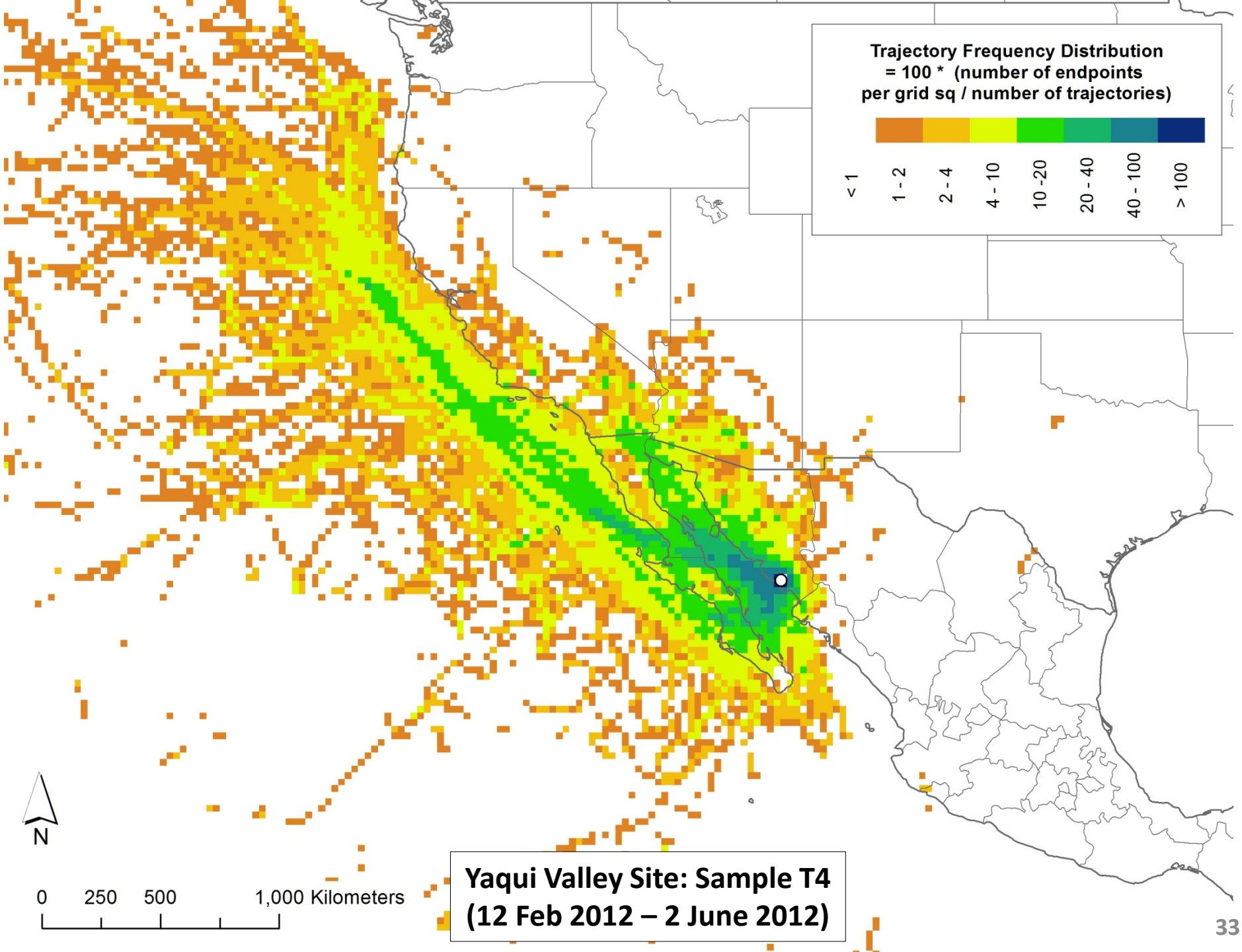
Yaqui Valley Site: Sample T1 (25 Mar 2011 – 1 July 2011)



Yaqui Valley Site: Sample T2
(1 July 2011 – 7 Oct 2011)

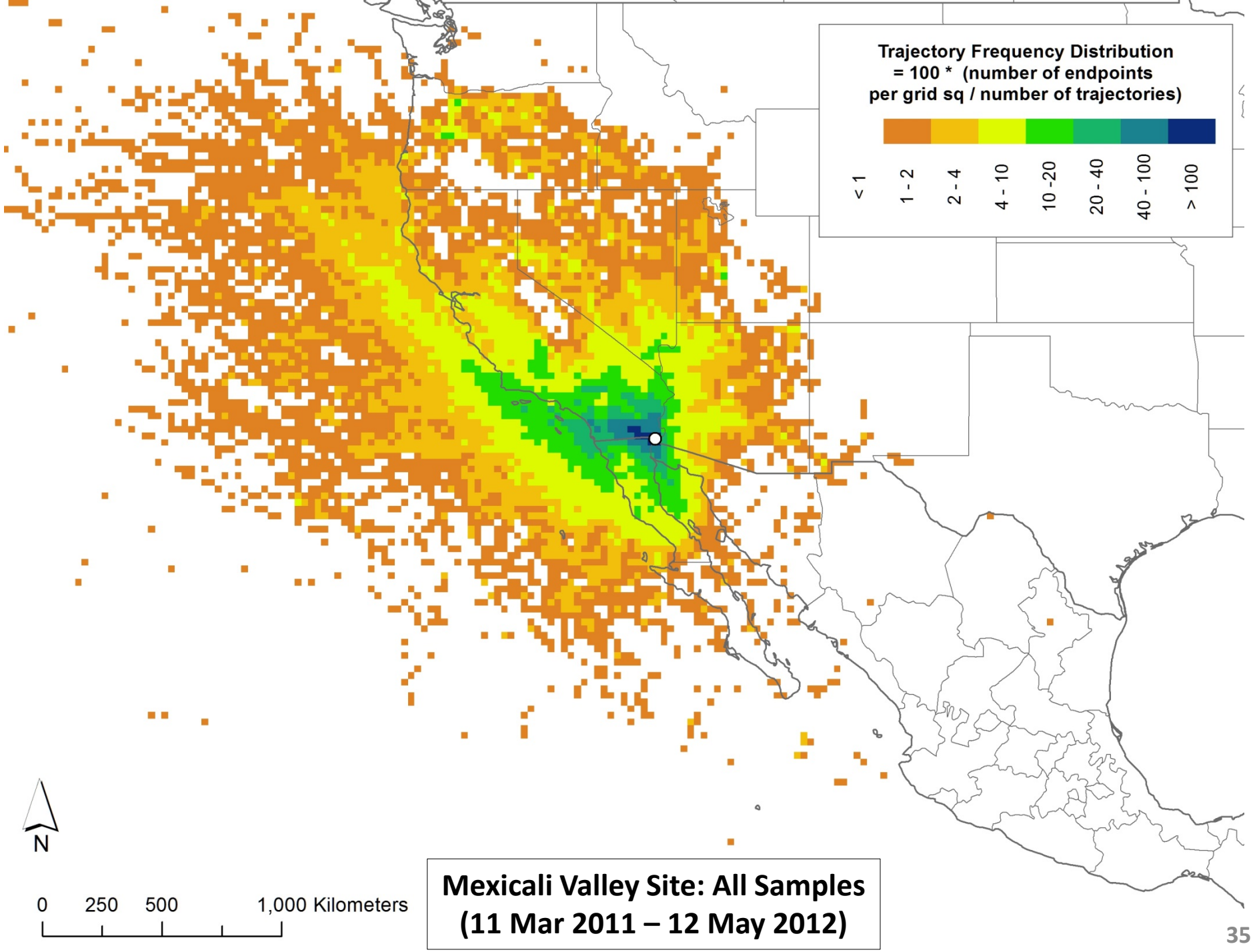


Yaqui Valley Site: Sample T3
(7 Oct 2011 – 12 Feb 2012)

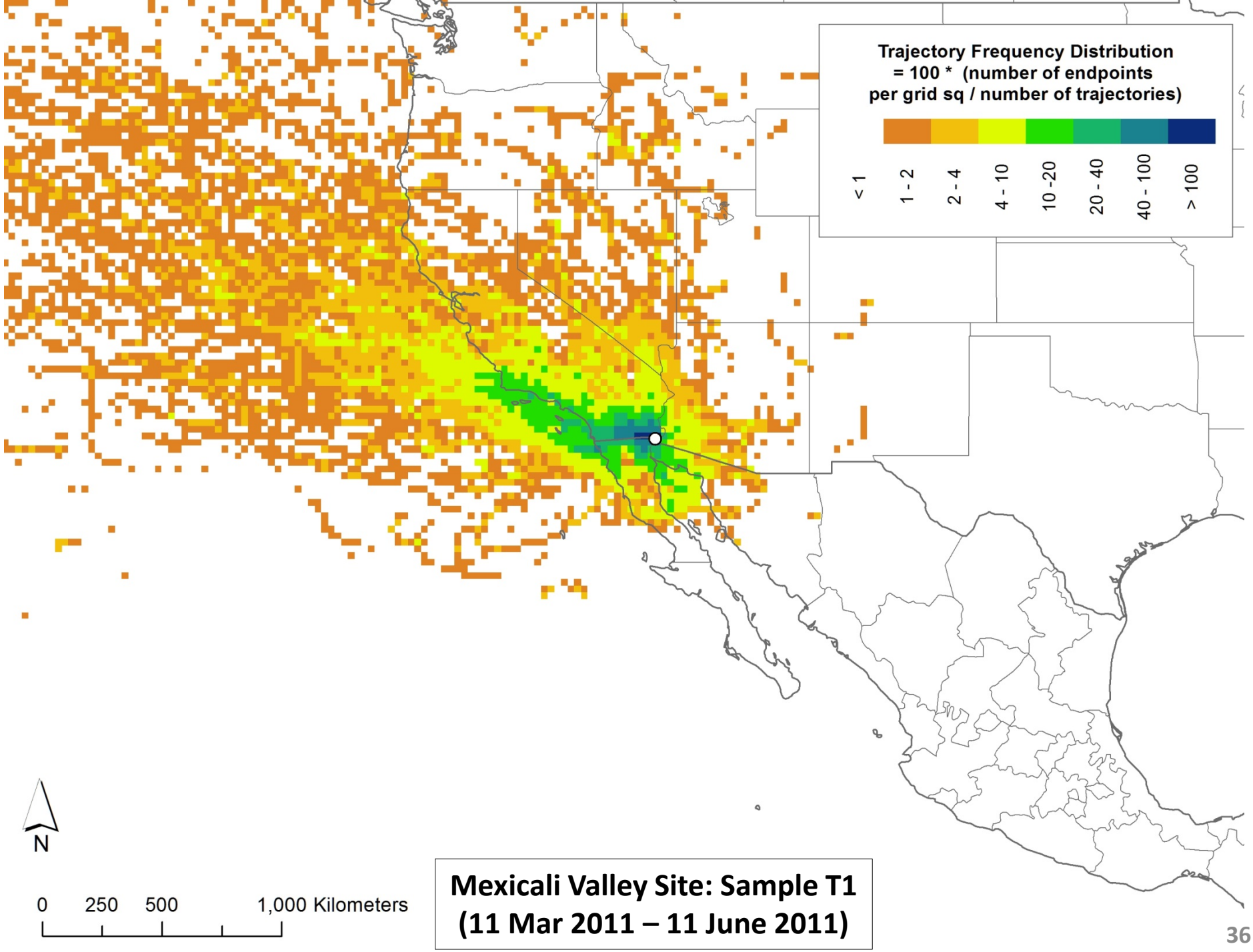


Yaqui Valley Site: Sample T4
(12 Feb 2012 – 2 June 2012)

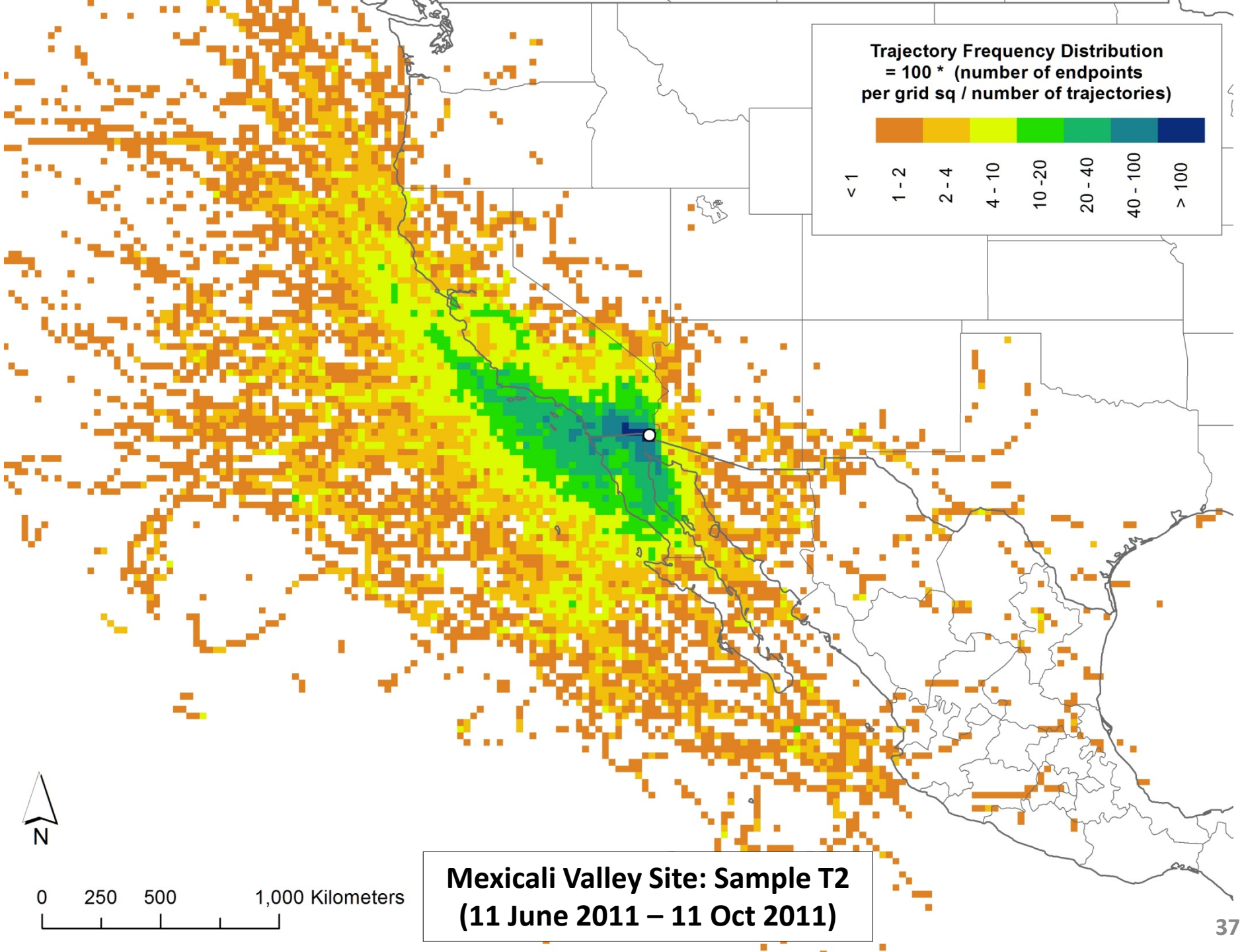
Comparable Graphics for the Mexicali Valley samples



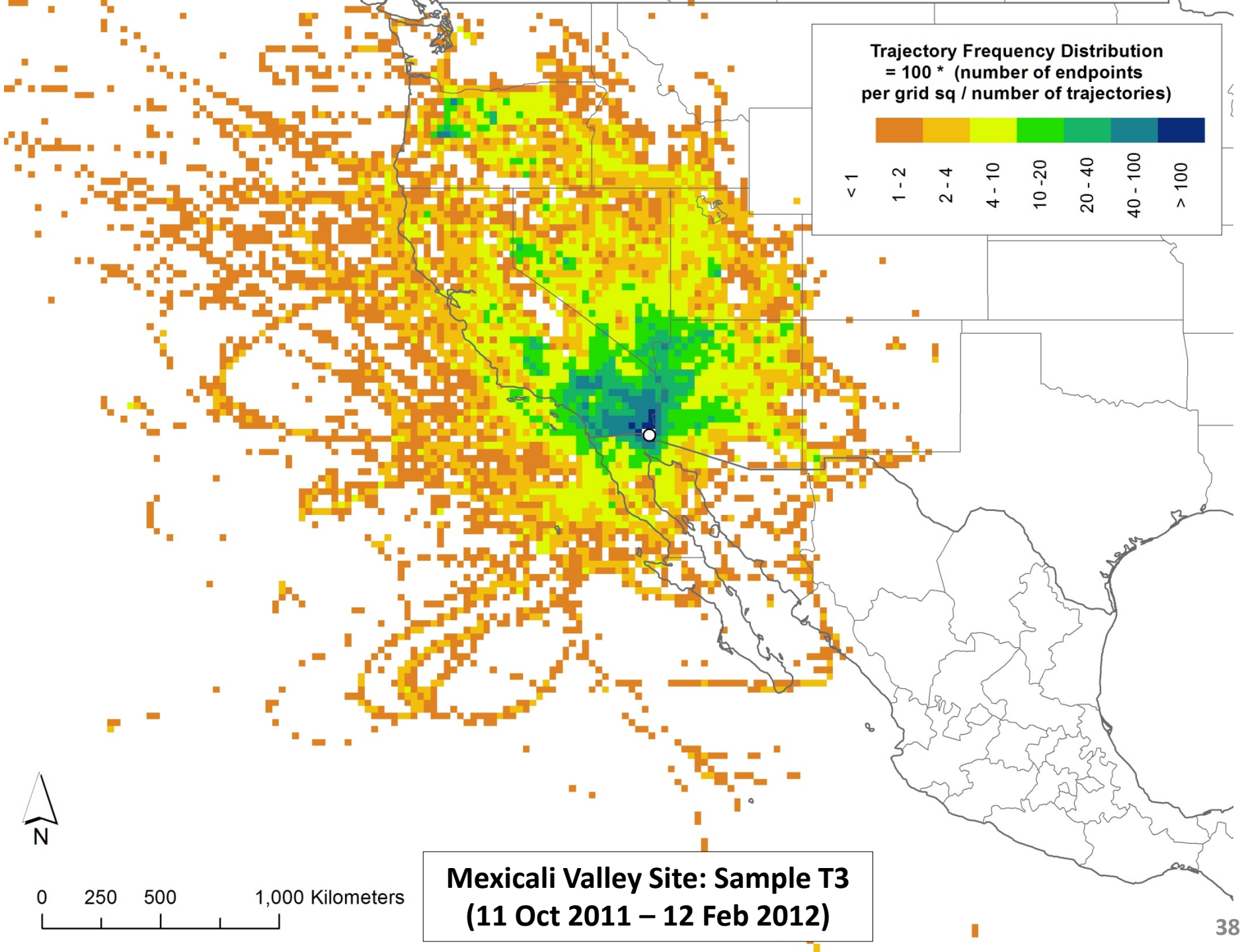
Mexicali Valley Site: All Samples (11 Mar 2011 – 12 May 2012)



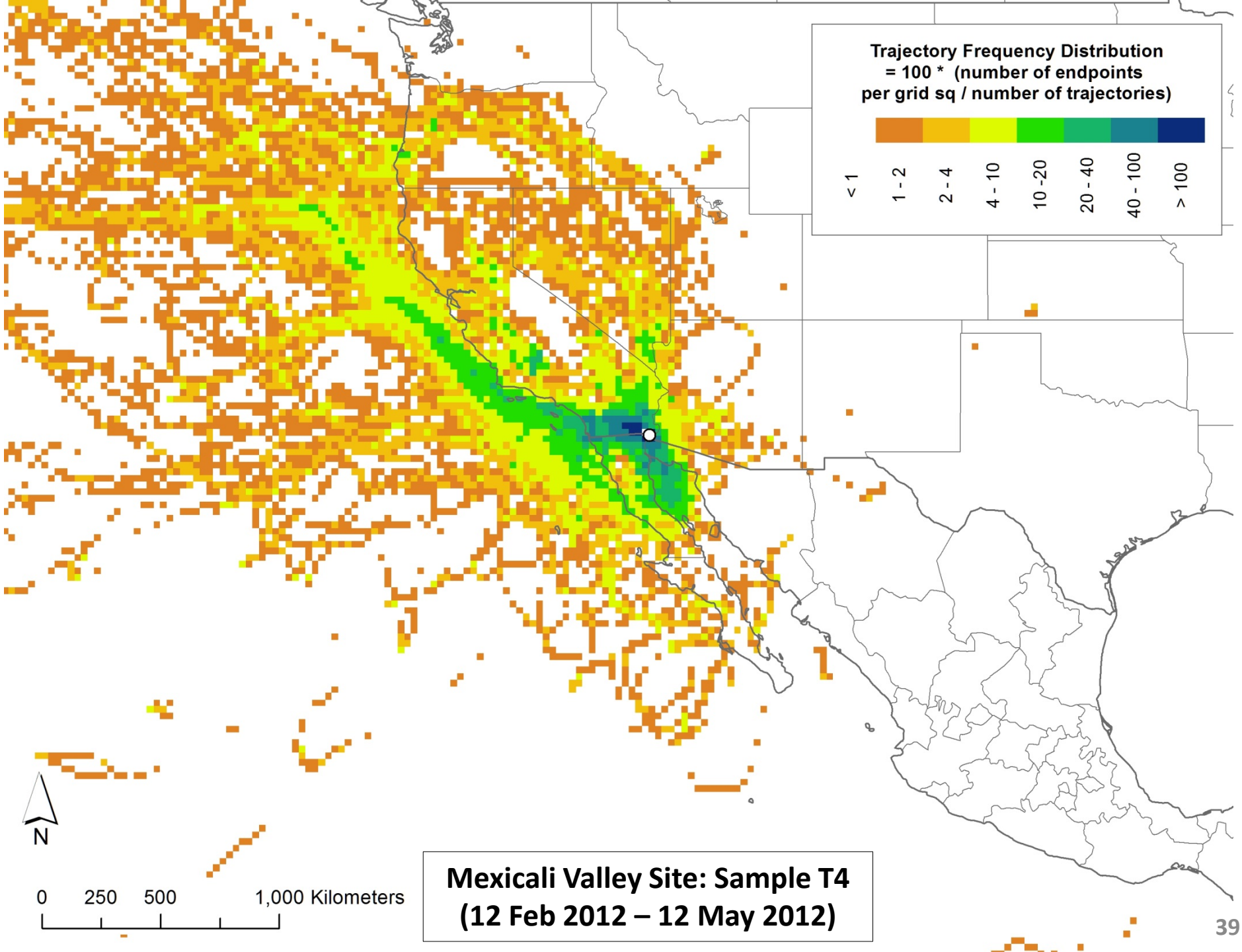
Mexicali Valley Site: Sample T1
(11 Mar 2011 – 11 June 2011)



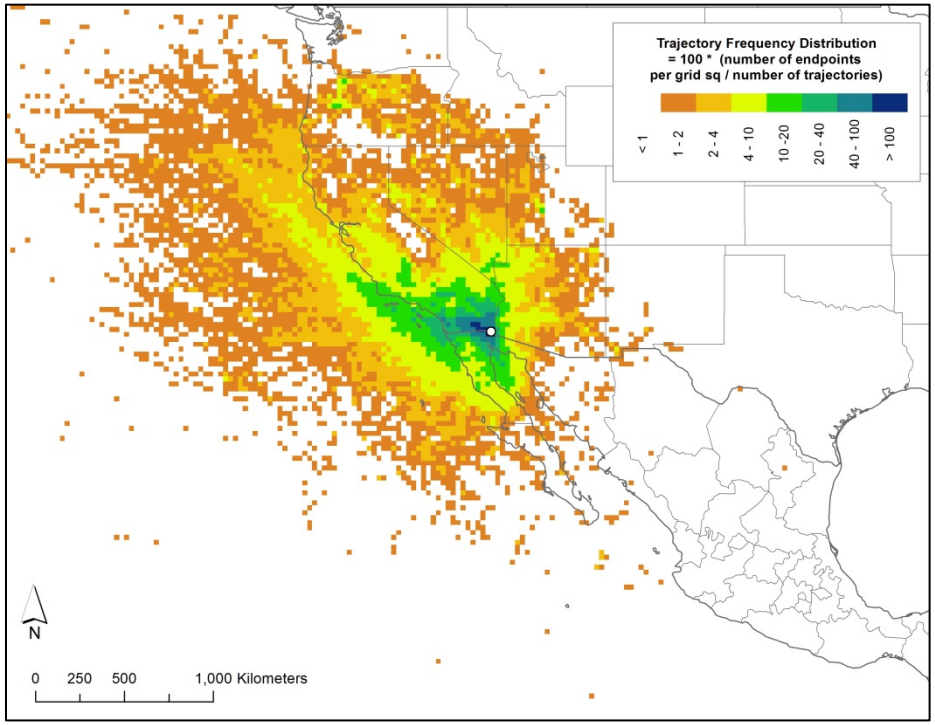
Mexicali Valley Site: Sample T2
(11 June 2011 – 11 Oct 2011)



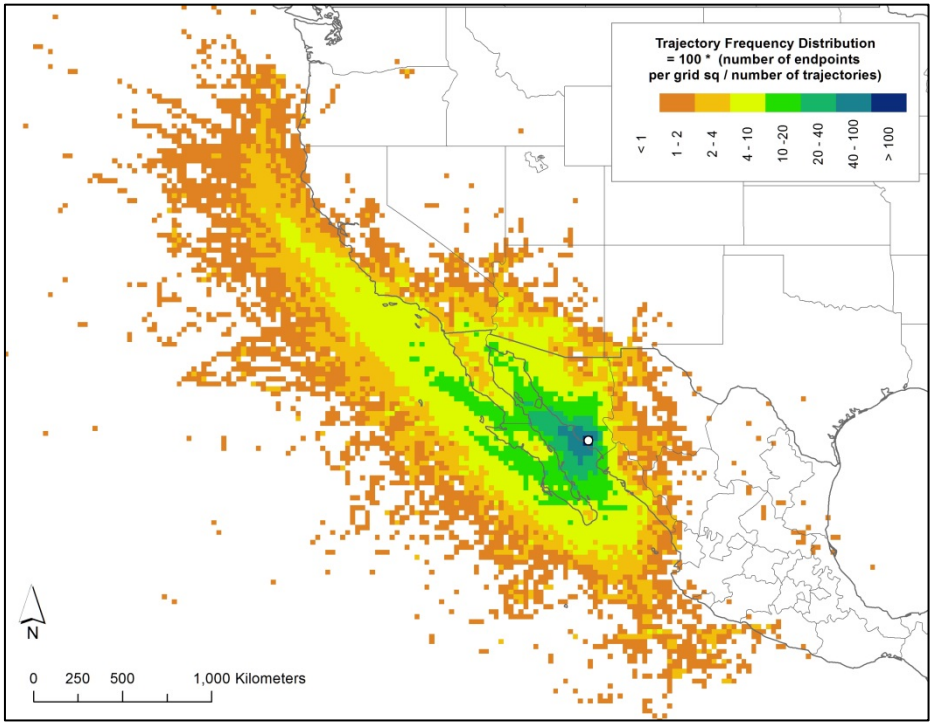
Mexicali Valley Site: Sample T3
(11 Oct 2011 – 12 Feb 2012)



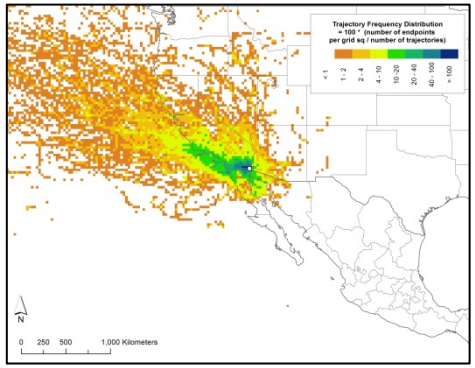
Mexicali Valley Site: Sample T4
(12 Feb 2012 – 12 May 2012)



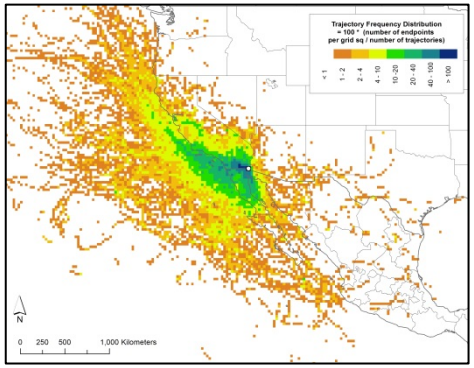
Mexicali Valley Site: All Samples
(11 Mar 2011 – 12 May 2012)



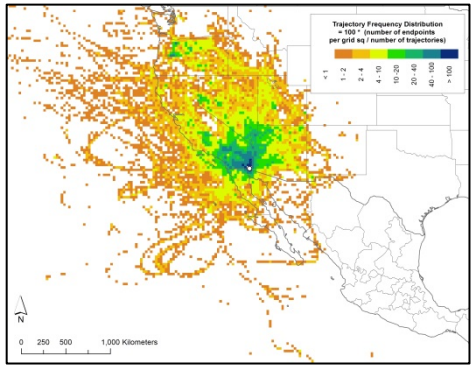
Yaqui Valley Site: All Samples
(25 Mar 2011 – 2 June 2012)



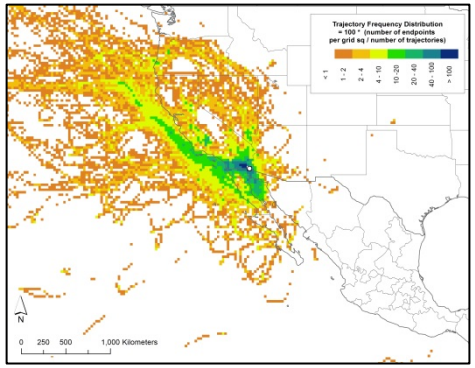
Mexicali Valley Site: Sample T1
(11 Mar 2011 – 11 June 2011)



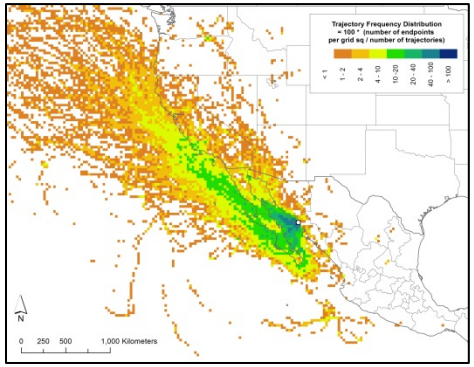
Mexicali Valley Site: Sample T2
(11 June 2011 – 11 Oct 2011)



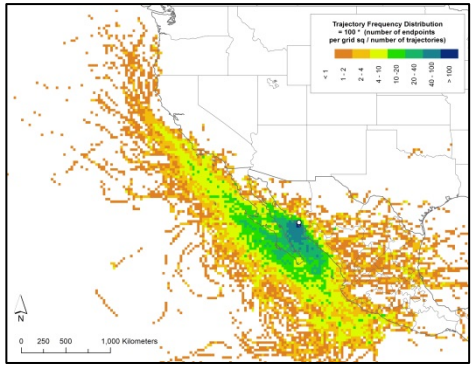
Mexicali Valley Site: Sample T3
(11 Oct 2011 – 12 Feb 2012)



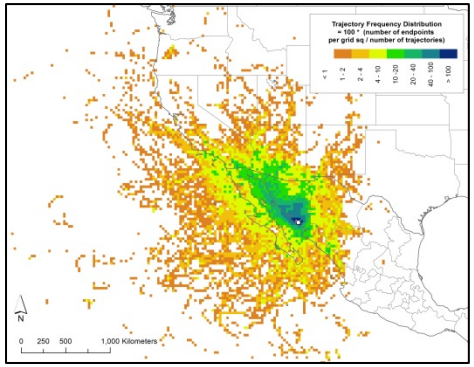
Mexicali Valley Site: Sample T4
(12 Feb 2012 – 12 May 2012)



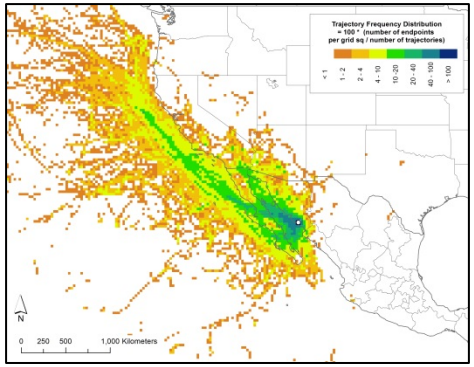
Yaqui Valley Site: Sample T1
(25 Mar 2011 – 1 July 2011)



Yaqui Valley Site: Sample T2
(1 July 2011 – 7 Oct 2011)

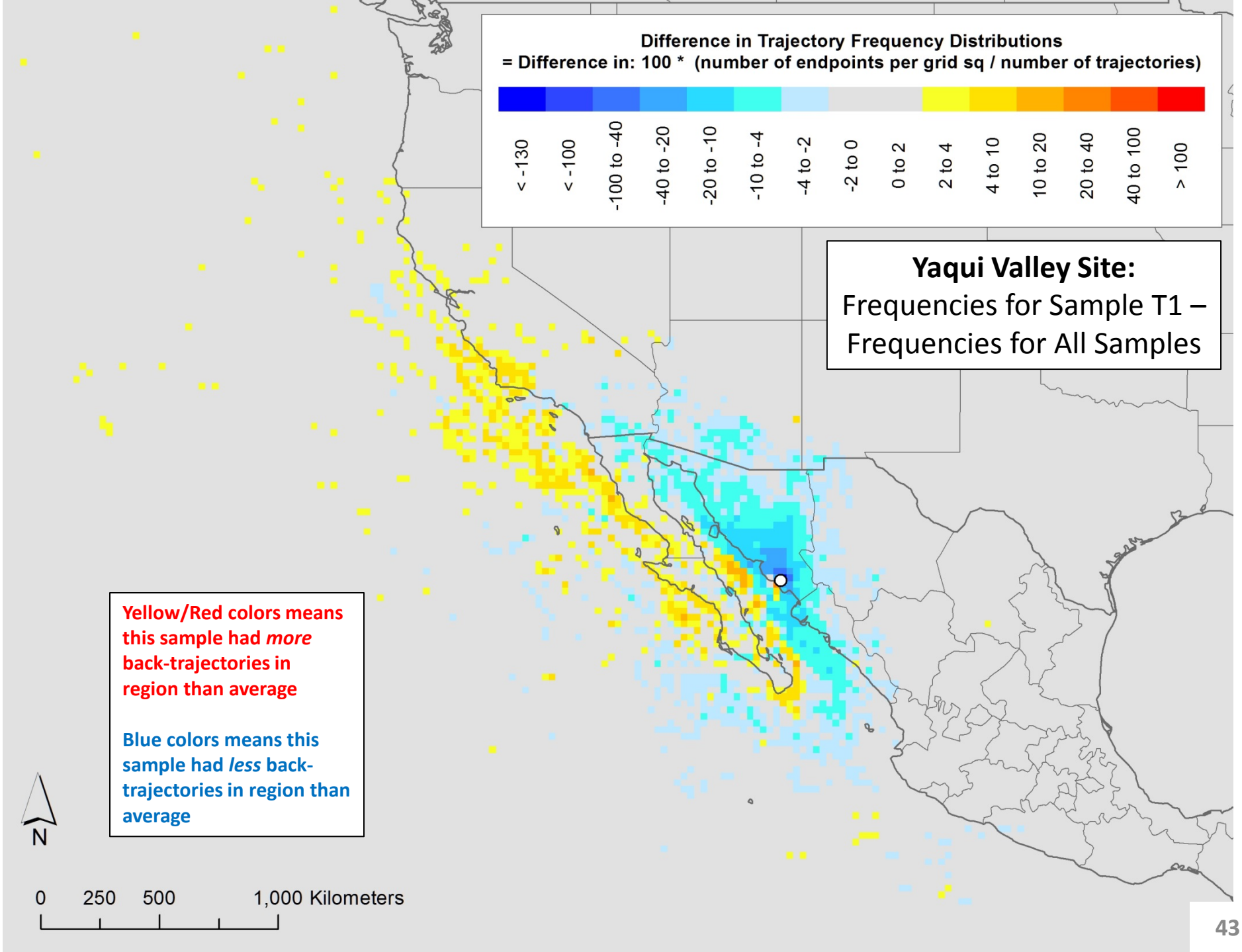


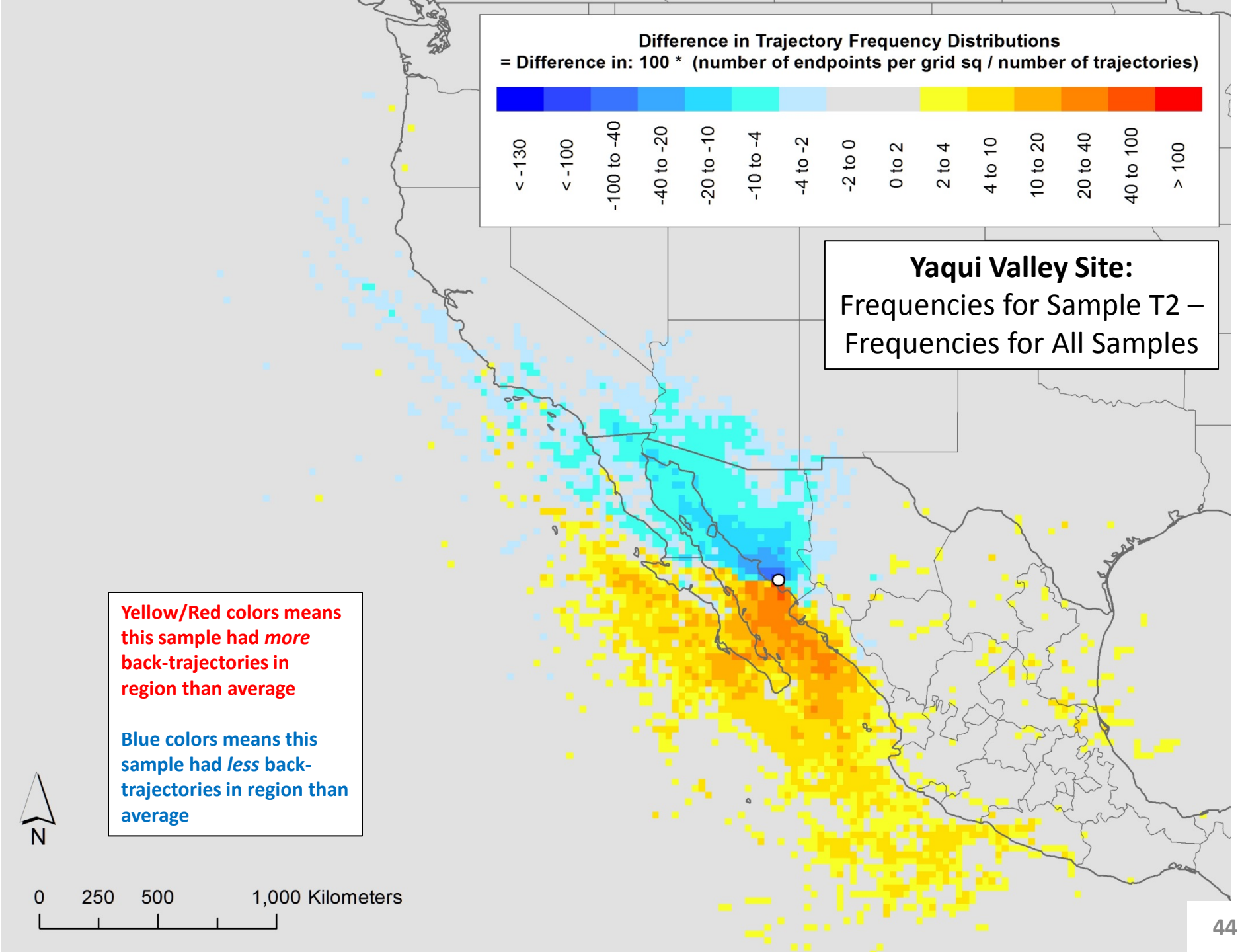
Yaqui Valley Site: Sample T3
(7 Oct 2011 – 12 Feb 2012)

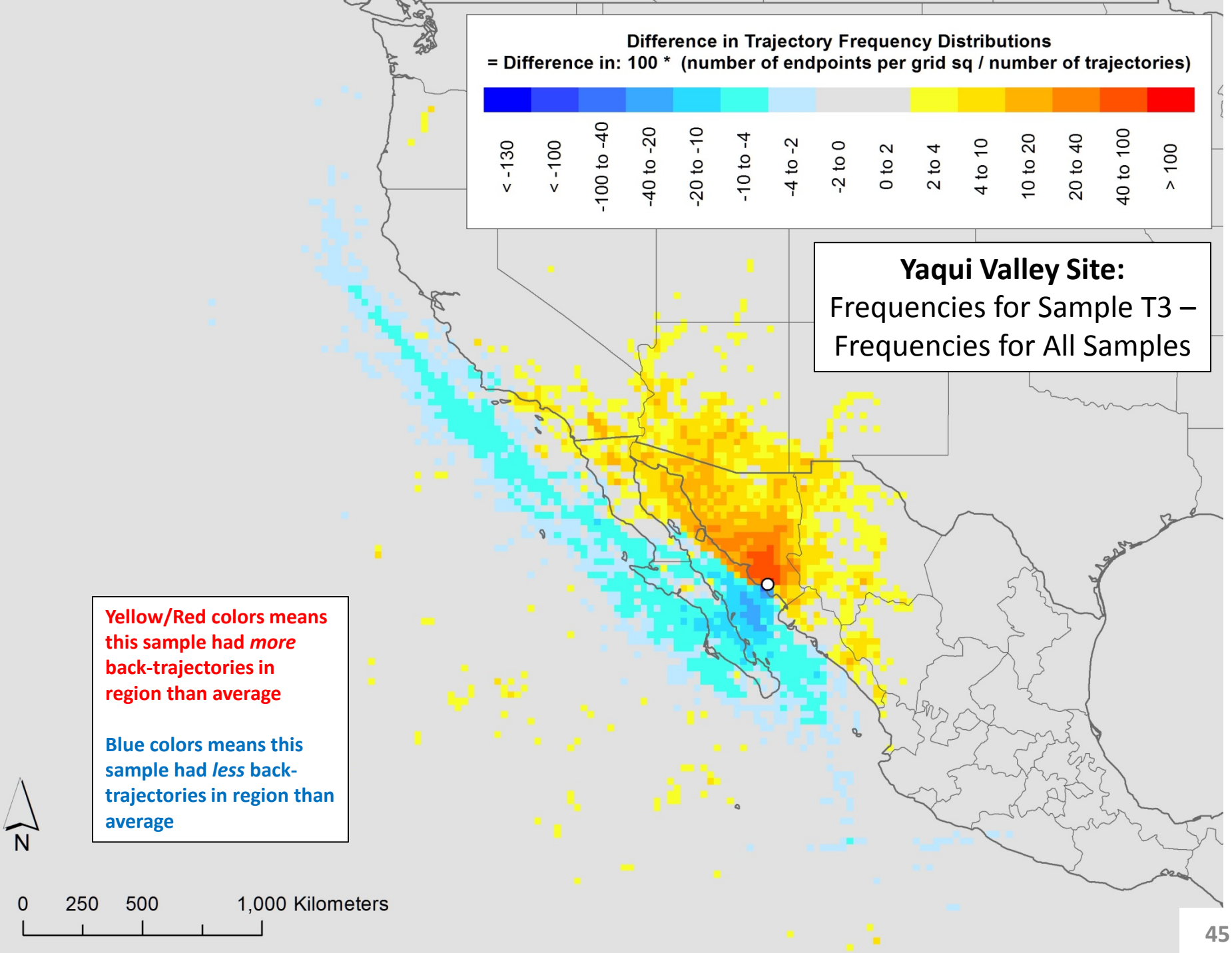


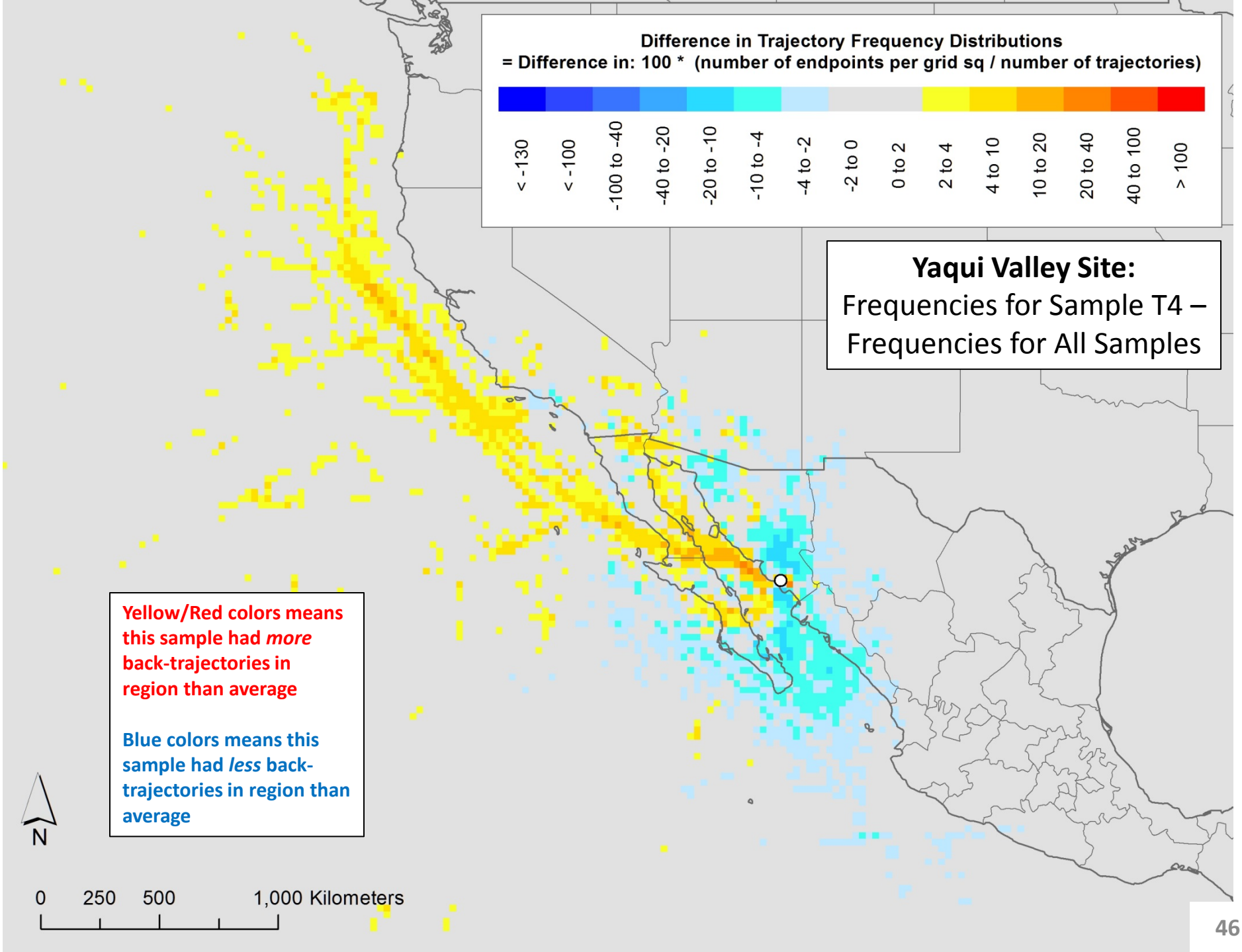
Yaqui Valley Site: Sample T4
(12 Feb 2012 – 2 June 2012)

Here’s another way to compare the frequency results

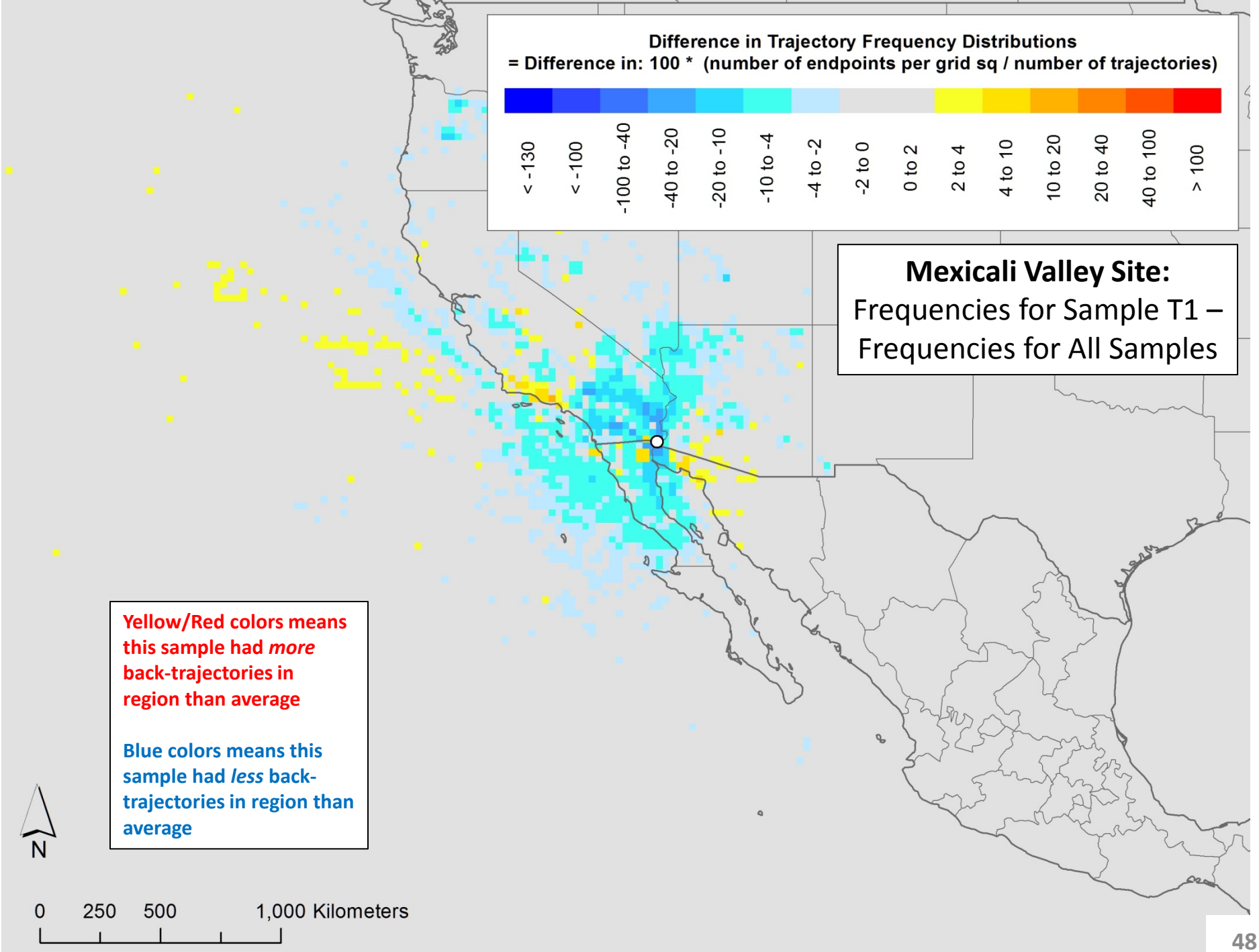

Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Yaqui Valley Site:
Frequencies for Sample T1 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers

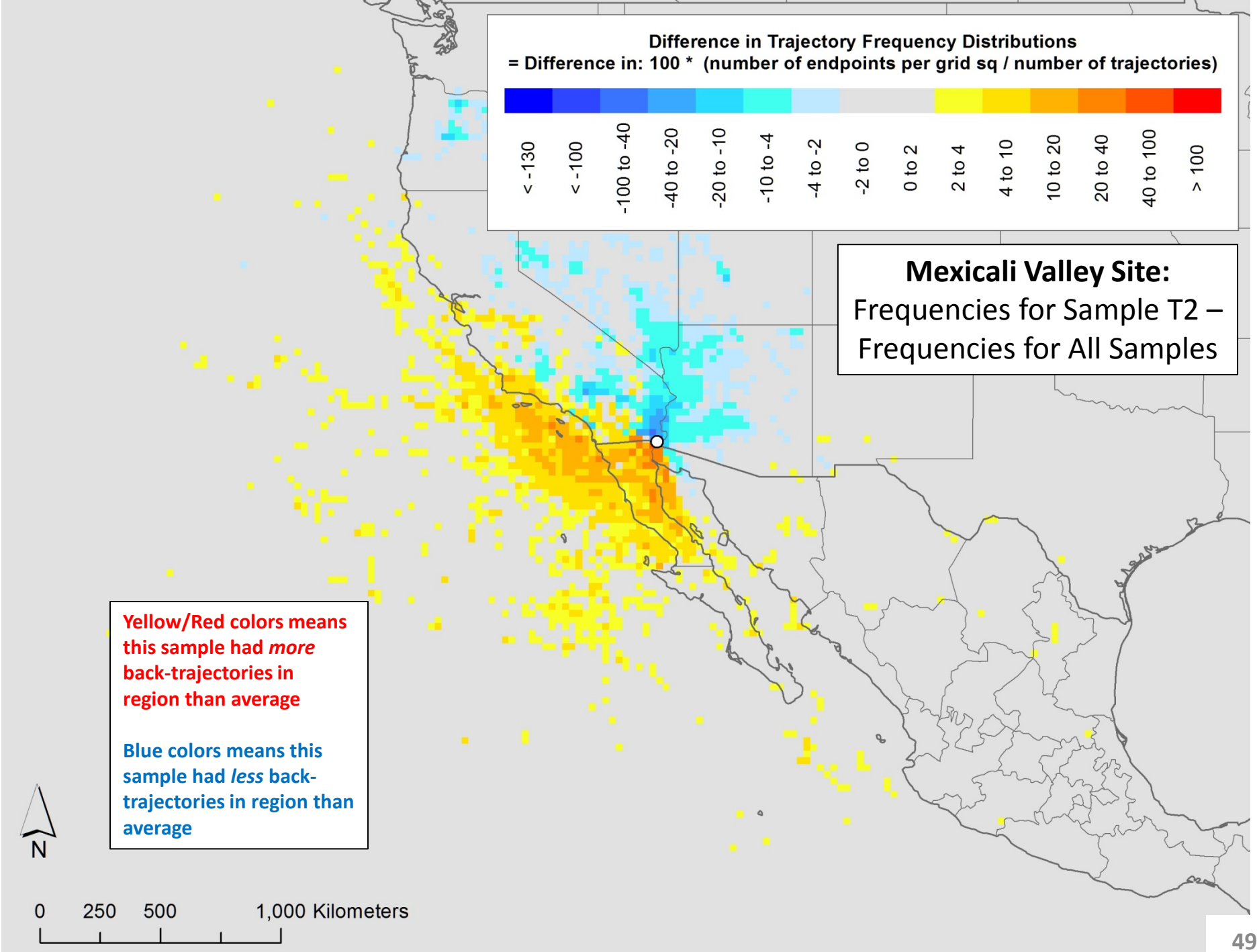

Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Yaqui Valley Site:
Frequencies for Sample T2 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers

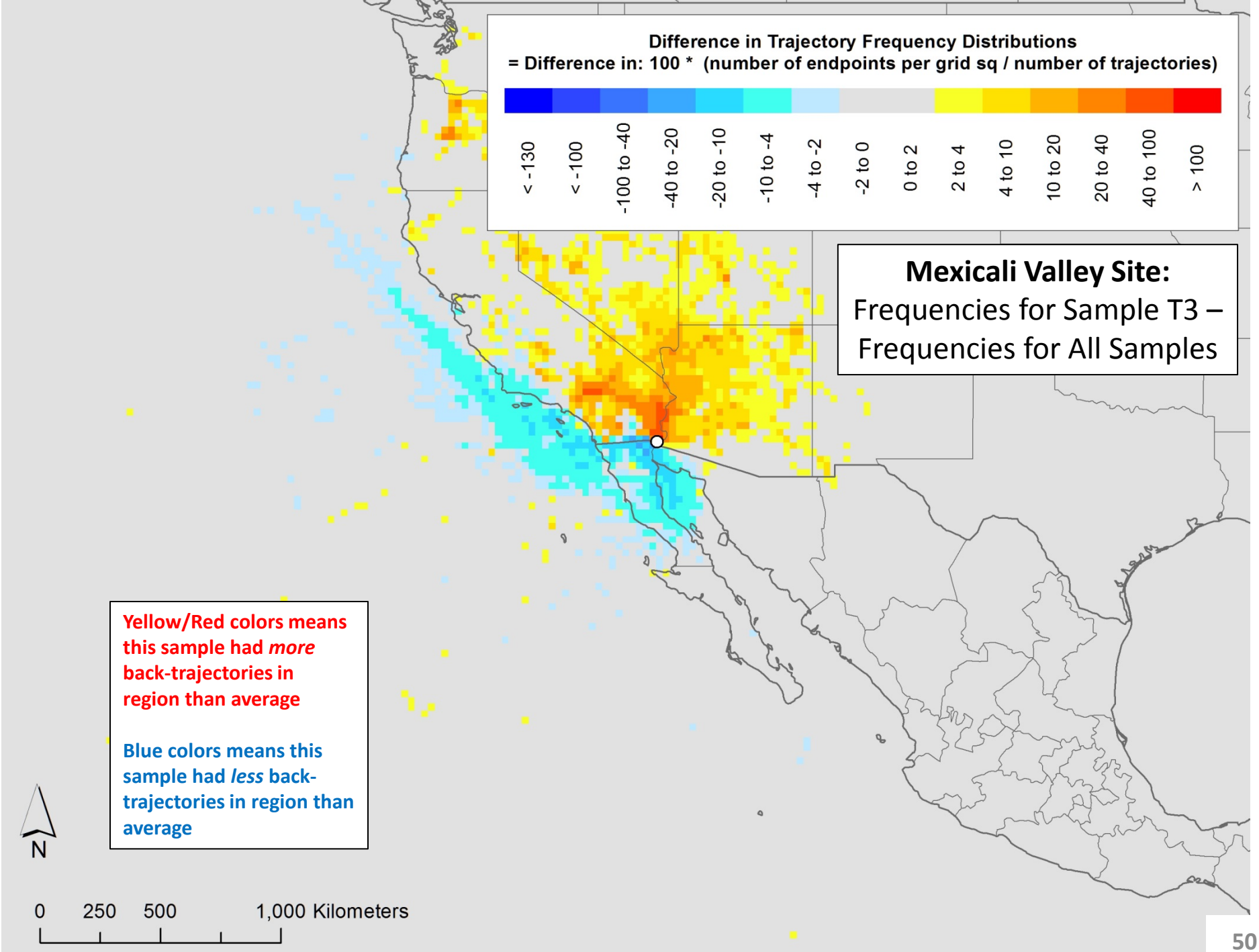

Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Yaqui Valley Site:
Frequencies for Sample T3 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers

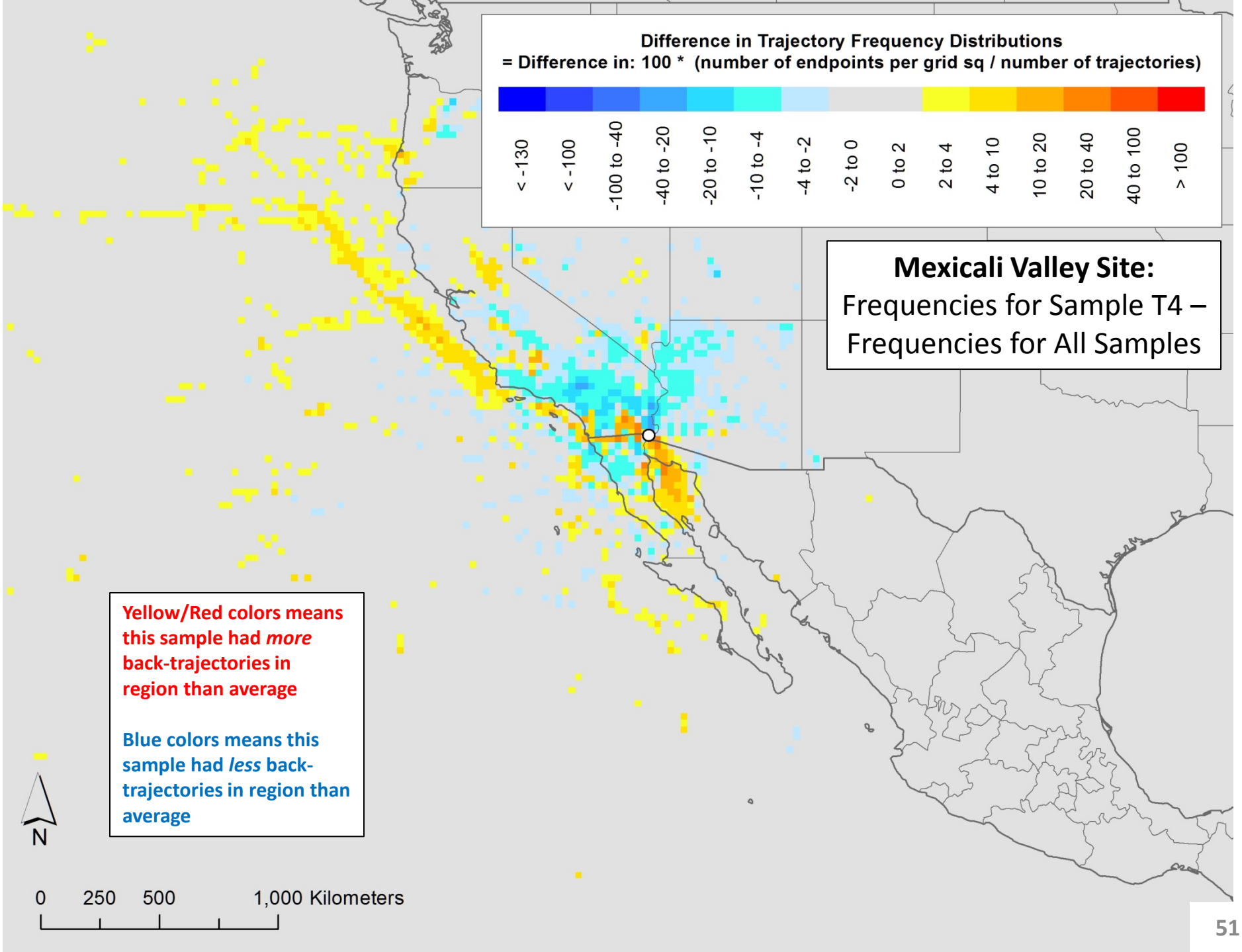

Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Yaqui Valley Site:
Frequencies for Sample T4 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers

Comparable Graphics for the Mexicali Site Samples

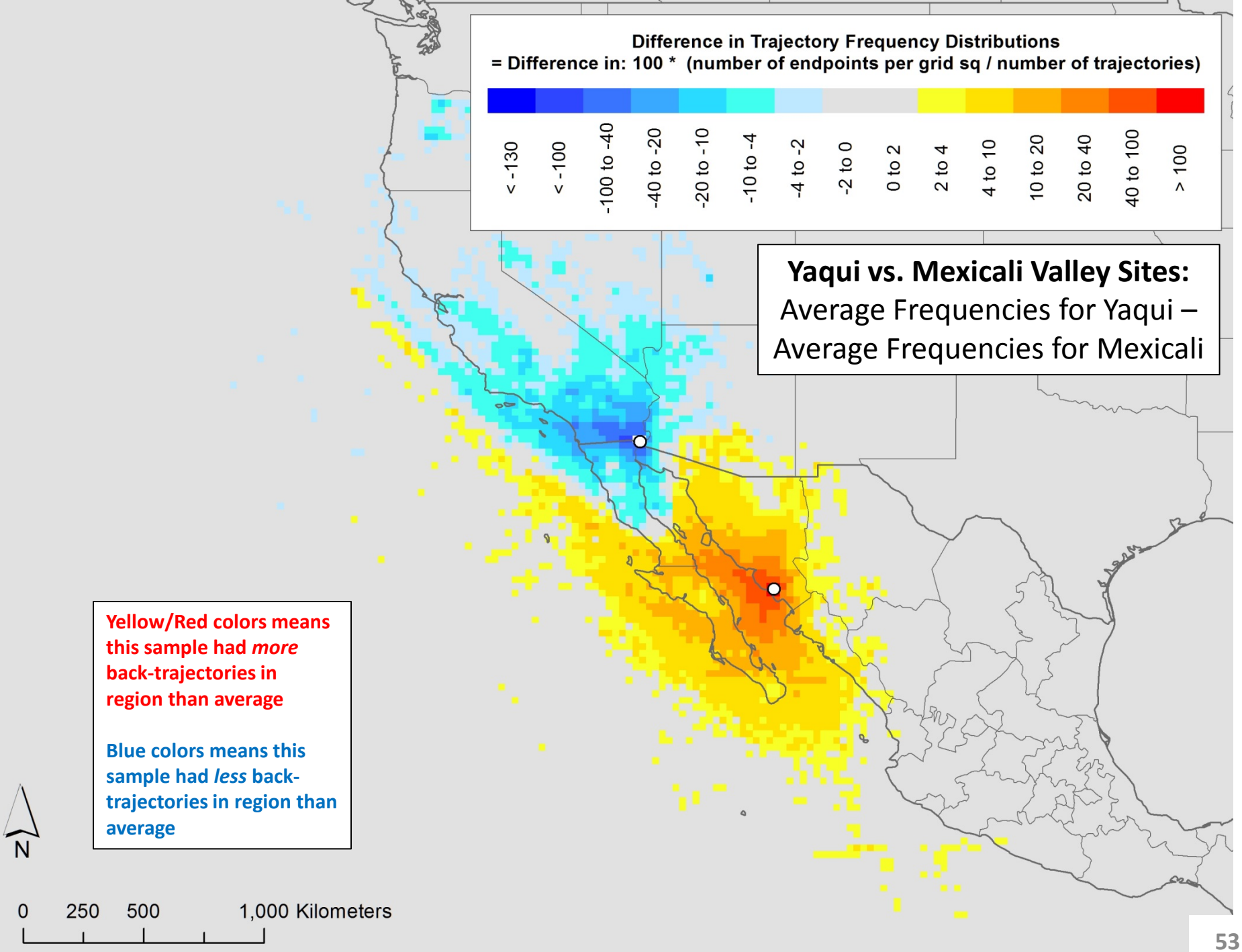

Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Mexicali Valley Site:
Frequencies for Sample T1 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers


Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Mexicali Valley Site:
Frequencies for Sample T2 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers


Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Mexicali Valley Site:
Frequencies for Sample T3 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers


Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Mexicali Valley Site:
Frequencies for Sample T4 – Frequencies for All Samples
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers

The next slide shows the difference between the entire set of Yaqui Valley back-trajectories and the entire set of Mexicali Valley site back-trajectories


Difference in Trajectory Frequency Distributions
= Difference in: 100 * (number of endpoints per grid sq / number of trajectories)
< -130
< -100
-100 to -40
-40 to -20
-20 to -10
-10 to -4
-4 to -2
-2 to 0
0 to 2
2 to 4
4 to 10
10 to 20
20 to 40
40 to 100
> 100
Yaqui vs. Mexicali Valley Sites:
Average Frequencies for Yaqui –
Average Frequencies for Mexicali
Yellow/Red colors means this sample had *more* back-trajectories in region than average
Blue colors means this sample had *less* back-trajectories in region than average
N
0 250 500 1,000 Kilometers

Outline of Procedure

1. Download/Install PC-HYSPLIT
2. Also might want to have a text editor that displays context highlighting (e.g., Notepad++)
3. Develop DOS "SET" batch file for running multiple trajectories
   - Basically, this writes CONTROL, SETUP.CFG, runs HYSPLIT, and creates output
   - Carefully consider what output frequency needed for endpoints. If frequency too low, results will not be suitable for fine-scale mapping
4. Test this "SET" batch file on a few trajectories
5. Excel-based "run-generator" to create "RUN" batch file for all needed trajectories
6. Run all of the trajectories
7. Determine if all trajectories finished
   - Each tdump ("TDP") file should be the same size
   - Each MSG file should be at least moderately sized
8. For any trajectories that did not finish, try to re-run them to get them to finish
   - Can examine MSG file
   - Can re-run to examine run-time error messages
9. Create "INFILES" for each sample for each site, and for all of the trajectories for each site
10. Develop DOS "SET" and "RUN" batch files for running TRAJFREQ
    - Select a starting grid size and extent
    - Can carry out runs for other grid sizes and extents with the same trajectories
11. Carry out TRAJFREQ runs for each sample, and for "ALL", for each site
12. Can adjust frequency ranges if desired for CONCPLOT output
13. Import frequency outputs into GIS for more advanced displays
14. Can run special program "FREQDIFF" for another way to display results