Application of Factor Analysis to the BRAVO Dataset

Prepared by

Texas Commission on Environmental Quality
Data Analysis Section
Fernando Mercado
Stuart Dattner
Bethany Georgoulias

Contents

I. Introduction

The potential decline of visibility at Big Bend National Park (BBNP) is an important problem the Big Bend Aerosol and Visibility Observational study (BRAVO) was designed to address. This paper focuses on the application of factor analysis to the BRAVO aerosol (PM2.5) dataset. Through factor analysis, it is possible to investigate the factor structure underlying the set of observed variables. This paper will also investigate seasonal, extreme and mean values, with respect to species related to factors identified through the factor analysis.

II. Data Limitations

During the BRAVO study, Caney Creek, AR and Wichita Mountains, OK (see figure 20) were the only sampling sites outside of Texas. Due to data filtering, these two sampling sites were not used (see II.a. - Data Preparation) in the analysis. Having no sampling sites outside of Texas, this data set cannot sufficiently address the issue of particulate matter being transported into Texas by winds. This is especially relevant because an analysis of maximum daily concentrations suggests that there were significant sources to the east and northeast of Texas that might have been missed by the monitoring network used in the BRAVO Study. Therefore the analysis must rely on other methods to estimate PM such as, back trajectories and modeling.

III.a. Data Preparation – All of Texas Sites

The data used in this analysis are 24-hour PM2.5 samples. The BRAVO aerosol data set included Quality Assurance (QA) flags and comment fields to filter the data appropriately, removing data values that indicate that the sample and/or the analysis of the sample may be unusable or unreliable. Statistical Analysis Software (SAS) factor analysis procedure (or any other procedure) requires that the data set matrix have no null values and a minimum number of records for each variable; Gorsuch[1] recommends five times the number of samples for each variable. This recommendation combined with filtering forced the omission of all the carbon and ion species. After the omission of certain species, the total numbers of records were 2,637 with 22 variables or species. The original data set contained 6,429 records and 180 variables or species. The following species were retained: Arsenic (As), Bromine (Br), Copper (Cu), Iron (Fe), Hydrogen (H), Total Particulate Mass (Mass), Aluminum (Al), Lead (Pb), Calcium (Ca), Chromium (Cr), Phosphorous (K), Manganese (Mn), Sodium (Na), Sulfur (S), Silicon (Si), Titanium (Ti), Vanadium (V), Rubidium (Rb), Selenium (Se), Strontium (Sr), Yttrium (Y) and Zinc (Zn) for the factor analysis; and this dataset is called the ‘elements only’ dataset which cover all Texas sites.

The following criteria were used during the filtering process. Each bullet represents conditions the variable concentrations must meet in order to be retained for analysis:

- All variables with concentration of zero or greater.
- All variables with flow rate through the filter between 21.3 and 24.3, corresponding to an aerodynamic particle 50% cut point of between 2.0 and 3.0 micrometers.

[1] Factor Analysis, Richard L. Gorsuch 1974

- All variables with concentrations equal or greater to minimum detection level (MDL).
- All variables with elapsed times between 1080 (18 hrs) and 1580 (26.3 hrs) minutes.
- All variables with filter status and machine status flags of NM and SC.
- All variables with a minimum of 110 entries.

After the above filtering, missing values were converted to zero to fill in the data matrix. This dataset includes most sampling sites in the BRAVO study (see Figure 20 for a map showing the location of the sampling sites).

III.b. Data Preparation – Big Bend Only

The adherence to commonly accepted data quality assurance techniques resulted in the omission of some of the more interesting species in the BRAVO dataset. To overcome this problem when examining the Big Bend area only, a second dataset with very relaxed filtering resulted in a truncated dataset with 270 observations and 37 variables that included only the Big Bend Sites. This dataset had a reduced number of observations but included more variables (species). This dataset contains the following species: Na, Al Si, S, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn As, Pb, Se Br, Rb, Sr, H, NO2, NO3, SO4, Na ion, NH4, K ion, Mg ion, Ca ion, O1, O2, O3, O4, OP, E1, E2, E3. The hope was that this truncated dataset would expose more detail regardless of the filtering practices. This dataset is called ‘elements, carbons and ions’ dataset and only covers Big Bend sites. For more information on filtering parameters see the appendices.

The data set called elements, carbons and ions was filtered with the following criteria:

- All variables have a minimum of 75 observations.
- All observations have concentration not equal to –99.

After the above filtering, missing values were converted to zero to fill in the data matrix. A scatter plot matrix was created for each of the datasets and inspected for leverage points and obvious outliers. No data points were found to have an unreasonable influence; therefore, no data points were omitted.

IV. Factor Analysis Procedure

Factor analysis is performed on a matrix of correlations, and the data should satisfy the assumptions for the Pearson correlation coefficient. One assumption is that there exists a linear relationship between all observed variables. This assumption has not been verified in the datasets. Secondly, each observed variable should be normally distributed. A histogram and a normal probability plot were examined for each variable, and none of the variables exhibited a normal distribution. Despite the deviation from the assumptions, the hope is that the factors will contain useful and meaningful information. Factor analysis was conducted in exploratory form.

The initial step in the factor analysis was to choose the appropriate number of factors to extract using squared multiple correlations (SMC) as prior communality estimates for each of the two datasets. Principle factor method was used to extract the factors, and this was followed by Varimax (orthogonal) rotation. SAS software was used to apply the factor analysis to each dataset and SAS provides scree plots for choosing the appropriate number of factors to retain. The resulting scree plot suggested possibly four meaningful factors; therefore, four factors were retained for the 'elements only' dataset. For the dataset named 'elements, carbons and ions', the scree plot suggested six meaningful factors, and so six factors were retained. More factors were retained than less at the risk of having unresolved factors as Cattell[2] suggests.

In interpreting the rotated factor pattern for both datasets, an item was said to load on a factor if the absolute value of the factor loading was greater than 0.30 for each dataset. A negative loading corresponds to sources being on opposite sides of the receptor. Using these criteria we have the following results (See Tables 1.a, 1.b, 2.a, 2.b).

Once factor loadings were acquired, the top 50 factor loadings for a given factor were contoured and displayed on a regional map. Higher factor loadings may indicate a potential source or sources for a particular factor (see Figure 1).

V. Rotated Factor Pattern Results

[2] Cattell, 1952, 1958, Rummel, 1970.

Rotated Factor Pattern For Elements Only Dataset
(State Wide)

	Factor1		Factor2		Factor3		Factor4	
Na	47	*	-1		6		-13	
Al	99	*	-4		-3		-5	
Si	99	*	-2		2		-3	
S	-4		93	*	4		5	
K	80	*	11		47	*	4	
Ca	55	*	-3		27		4	
Ti	92	*	19		-6		-3	
V	56	*	-22		10		-3	
Cr	5		0		19		5	
Mn	71	*	10		10		7	
Fe	99	*	0		1		1	
Cu	-3		7		7		54	*
Zn	-4		48	*	43	*	51	*
As	11		30	*	42	*	4	
Pb	-1		24		17		72	*
Se	-12		58	*	10		24	
Br	-8		40	*	69	*	14	
Rb	92	*	-3		-1		-4	
Sr	93	*	-2		7		-2	
Y	36	*	16		-5		20	
Mass	39	*	82	*	22		12	
H	-3		93	*	20		15	

Printed values are multiplied by 100 and rounded to the nearest integer. Values greater than 0.3 are flagged by an '*'.

Table 1.a

Summary Table 1.b – Significant Factor Loadings

Factor 1	Factor 2	Factor 3	Factor 4
Na, Al, Si, K, Ca, Ti, V, Mn, Fe, Rb, Sr, Y, Mass	S, Zn, As, Se, Br, Mass, H	K, Zn, As, Br	Cu, Zn, Pb

Table 1.b: Summary to Table1.a - Rotated Factor Pattern (state wide).

Rotated Factor Pattern for Elements, Ions and Carbons Dataset

(Big Bend Only)

	Factor1		Factor2		Factor3		Factor4		Factor5	
NA	24		13		10		22		38	*
AL	97	*	-2		-5		11		8	
SI	98	*	0		-1		10		7	
S	1		96	*	12		2		5	
K	92	*	8		28		9		9	
CA	90	*	1		29		7		5	
TI	91	*	25		-10		11		0	
V	42	*	-12		9		-5		-44	*
CR	-13		4		-8		-4		15	
MN	71	*	14		8		4		-15	
FE	98	*	0		-1		11		7	
NI	-1		5		-3		1		33	
CU	-12		14		29		-19		-13	
ZN	14		31		73	*	-13		-22	
AS	12		32		17		-10		17	
PB	18		15		61	*	-5		-31	
SE	-5		68	*	3		-3		19	
BR	12		48	*	65	*	0		17	
RB	75	*	-4		2		6		-8	
SR	83	*	1		3		14		35	
H	11		95	*	18		3		-4	
NO2	-7		3		20		-11		10	
NO3	50	*	7		31		28		62	*
SO4	-1		93	*	10		14		8	
NA ION	48	*	2		0		25		64	*
NH4	-5		92	*	16		13		4	
K ION	23		23		52	*	14		31	
MG ION	64	*	-4		0		4		3	
CA ION	58	*	-6		42	*	22		16	
O1	46	*	-3		-11		34		-29	
O2	13		61	*	18		64	*	0	
O3	16		21		50	*	45	*	1	
O4	-5		61	*	53	*	31		8	

Rotated Factor Pattern for Elements, Ions and Carbons Dataset

(Big Bend Only)

		Factor1		Factor2		Factor3		Factor4		Factor5	
	OP	19		63	*	17		62	*	-1	
	E1	-16		72	*	39	*	43	*	3	
	E2	23		27		-10		74	*	18	
	E3	16		-4		-25		63	*	19	

Printed values are multiplied by 100 and rounded to the nearest integer.
Values greater than 0.356783 are flagged by an '*'.

Table 2.a Continued: Results from six factors retained.

Summary Table 2.b – Significant Factor Loadings
(Big Bend Area Only)

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Al, Si, K, Ca, Ti, V, Mn, Fe, Rb, Sr, NO3, Na ion, Mg ion, Ca ion, O1	H, S, Zn, As, Se, Br, SO4, NH4, O2, O4, OP, E1	Zn, Pb, Br, NO3, K ion, Ca ion, O3, O4, E1	O1, O2, O3, OP, E1, E2, E3	Na, V (neg.), NO3, Na ion

Table 2.b: Summary to Table2.a - Rotated Factor Pattern (Big Bend only).

VI. Interpretation of Factors

Elements Only Dataset, Relevant State Wide

Summary Table 1.b suggests that Factor 1 is related to crustal or soil elements. Factor 1 accounted for approximately 57 percent (7.92 out of 13.81 of the total variance, final communality estimates). Factor 1 top 50 loading scores had high loadings at Fort Lancaster and a region that stretches from Still House to Laguna Atacosta (see Figure 1).

Factor 2 (Table 1.b) suggested a relationship to coal because of the significant loadings from sulfur (S), hydrogen (H) and Selenium (Se). Selenium is a tracer for coal burning. Factor 2 accounted for 30 percent of the total variance (4.82 of the variance). Factor 2, top 10 factor loadings were primarily in the northeast part of Texas where there is a higher density of sulfur dioxide (SO_2) emitters, and possible coal combustion sources (see Figure 2 and 3). There are three SO_2 point sources north of the Big Bend area, one at Monahans Sand Hills, one at Fort Lancaster, and one at Mc Donald observatory which emit approximately 0.5, .01 and less than .003 ton per day of SO_2 respectively. These are the closest three SO_2 point sources near Big Bend, listed in the Point Source Data Base

(PSDB) for Texas[3]. Two other known sources of coal combustion are Carbon I and Carbon II in Mexico. The exact amounts of emissions from these sources are not known.

Factor 3 was an undetermined factor and accounted for approximately 7 percent or 0.98 of the total variance (Figure 4). The combination of bromine (Br), potassium (K), zinc (Zn) and arsenic (As) are not completely unintelligible. Possible sources are the petroleum industry, which uses bromine to control algae in cooling towers. Bromine is also used in the production of oil and gas well completion drilling fluids, photographic film and paper, and present in saltwater and other processes. Possible arsenic sources are the lumber and agricultural industry in which arsenic is used as a wood preservative. Potassium can be associated with the vegetation burning. These industries and processes occur near the San Bernard site and may be possible sources associated with factor 3. The proximity of the San Bernard site to three point sources (a Reliant Energy industry, photographic business and a chemical wholesaler) may have affected the ability to sample representative regional concentrations. This factor is not well understood and needs to be examined more closely. If the number of factors are reduced to three, the species that comprise factor 3 are distributed among the other factors somewhat evenly.
Factor 4 suggested a smelter influence because of the combination of copper (Cu), zinc (Zn) and lead (Pb), and accounted for 4.6 or 0.64 percent of the total variance (see Figure 5). The spatial distributions of factor 4 factor loadings to known locations of smelters and metal foundries, somewhat reflects the general location of the metal related processes (Figure 6) and influence on factor 4. Cu, Zn, and Pb can also be emitted from metal fabricators or municipal incinerators.

Elements, Carbons and Ions Dataset, Relevant to Big Bend Only

Summary Table 2.b suggests that Factor 1 was most likely associated with soil, with the most prominent factor loadings (>0.60) coming from elements like aluminum (Al), silicon (Si), potassium (K), calcium (Ca), titanium (Ti), manganese (Mn), iron (Fe), rubidium (Rb), strontium (Sr), and the magnesium ion (Mg^{+2}). Other variables with significant loadings on this factor were vanadium (V), nitrate (NO_3), the sodium ion (Na^+), and one of the organic carbon groups, O1. F_1 accounted for 33 percent of the total variance. Factor two (F_2), explaining 23 percent of the variance, was sulfate-related, with the heaviest loadings (>0.90) coming from sulfur (S), hydrogen (H), sulfate (SO_4), and ammonia (NH_4). Other important variables loading on this factor were selenium (Se), bromine (Br), organic carbon groups O2 and O4, and elemental carbon group E1. Selenium suggests this second factor could be associated with coal combustion, a primary source of sulfates[4]. Factor three (F_3) saw the largest factor loadings (>0.60) from zinc (Zn), lead (Pb), and Br. The K^+ and Ca^{+2} ions, as well as O3, O4, and E1, were also significant. Zinc and lead often indicate smelter influence[5]. F_3 explained 11 percent of the variance. Organic and elemental carbon groups O2, O3, OP, E1, E2, and E3 had significant loadings on factor four (F_4), while factor five (F_5) was associated with Na,

[3] TCEQ Point Source Database, 2000.
[4] Chow et al., 2002.
[5] Gebhart and Malm, 1997.

Na^+, and NO_3. These factors accounted for 10 and 7 percent of the variance, respectively. The last factor is consistent with other BRAVO findings that suggest that nitrates are present as sodium nitrate, rather than ammonium nitrate[6]. Interestingly, V had a strong negative loading on F_5 (-0.44), which meant significant vanadium concentrations were not present with the other species that loaded on this factor. Since the dataset containing elements, carbons and ions contained data from the Big Bend area only, spatial plots are not appropriate.

VII. Maximum Daily Concentrations

The frequency where the maximum concentration occurs for each day of the study were calculated for all species associated with the previously defined factors. The result was that the northeast Texas sites recorded the daily maximum concentration, of the entire domain (all BRAVO sampling sites) more often. This was also true if one subdivided the complete study period into two sub periods, defined as the following: The early part of the study defined as June1 – August 31 and the later part of the study defined as September 1 – October 31. Because the northeast sites recorded more of the daily maximum concentrations there are possibly strong sources nearby or influences by transport of pollutants from the East and Northeast of Texas. This result may also suggest that transport plumes tend to affect these northeast sites more often. In the early part of the study the northeast sites (Center and Wright Patman) tended to record approximately 40 percent of the daily maximum concentrations. In the late part of the study the northeast sites (Center and Big Thicket) recorded more daily maximum concentrations but the percentage dropped to about 25 percent. This suggests that there are seasonal influences, which tend to favor higher extreme values in the early part of the study. See figures 7 and 8.

VIII. Spatial Average Concentrations

Spatial average concentrations were calculated by taking the mean of all observations at each site and then contouring these averages for the two time periods as described above. All of the average concentrations were composed of at least 20 samples per site. For factor 1, the soil related species concentrations tended to be higher during the early part of the study and the higher average concentrations tended to cluster along the lower Rio Grande. The late part of the study shows much lower concentrations for most of Texas but still higher along the Rio Grande. The higher concentrations of soil species along the Rio Grande are a sharp contrast to other factors, which show higher concentrations in the northeast of Texas. Big Bend sites' concentrations were relatively low and did not change noticeably. See figures 9 – 11.

Factor 2 coal burning related species tended to be higher in concentrations during the early part of the study and were located in the northeast part of Texas. In the late part of the study the higher concentrations were more uniformly allocated and were distributed in the northeast and southeast part of Texas. There was a noticeable increase in the Langtry and Amistad region but not in Big Bend. See figures 12 –14.

[6] Hand et al., 2001.

Factor 3 (the unresolved factor associated with Br, As, K and Zn) species average concentrations remained constant throughout the study. The higher average concentrations were clustered in the east part of Texas. Big Bend concentrations did not dramatically change. See figures 15 and 16.

Factor 4 smelting associated species tended to be higher in the northeast part of Texas for both the early and late part of the study. In the second part of the study, noticeable average concentrations increases were recorded in west Texas. See figures 17 – 19.

IX. Conclusion

Factor analysis identified three factors in both datasets (Elements Only and Elements, Carbons and Ions): a soil related factor, a coal related factor and a smelter related factor. Finding these factors in both analyses indicates some robustness. This also indicates that there are common factors between Big Bend and Texas. Analysis of daily maximum concentration suggests that there is a strong influence either by nearby sources or transport, where the influence is the strongest in northeast Texas. Spatial average concentrations tended to be higher in the northeast part of Texas for most species. Many of the soil related species average concentrations tended to be higher in the first half of the study and clustered along the lower Rio Grande. Coal combustion factor related species concentrations were higher and more localized in northeast Texas during the first part of the study. The second part of the study a showed less localized pattern and lessened concentration over all east and central Texas. West Texas concentrations remained lower compared to east Texas for both parts of the study period. Smelting related factors were noticeable in the northeast and far west Texas. No one species was notably high at Big Bend compared to the rest of Texas.

The analysis done here suggests that there may be significant issues of aerosol transport from Mexico and other parts of the Continental United States that remain to be answered. The BRAVO study would benefit from additional aerosol monitoring in Mexico and to the East and Northeast of Texas in sufficient quantity and resolution to detect and quantify transport from those areas. The high concentrations in Northeast Texas may be due to transport from other states, but without monitors outside of Texas it will be difficult to recognize outside influence. Finally, because of the proximity of Big Bend to Mexico, monitoring in Mexico will be extremely useful in determining with greater accuracy the influence of Mexican emissions on Big Bend.

X. Figures

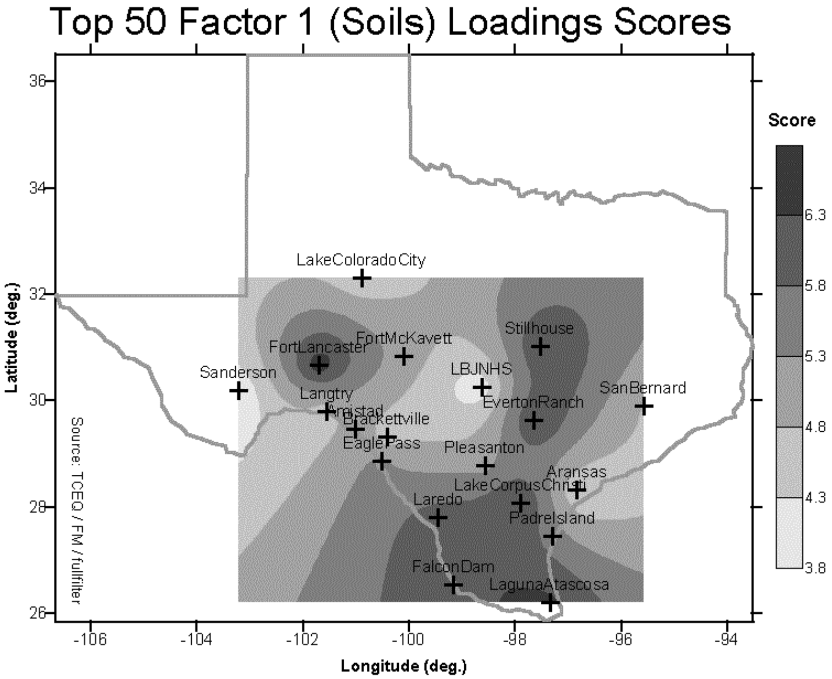


Figure 1- Demonstrates the location of the Top 50 loadings related to soil.

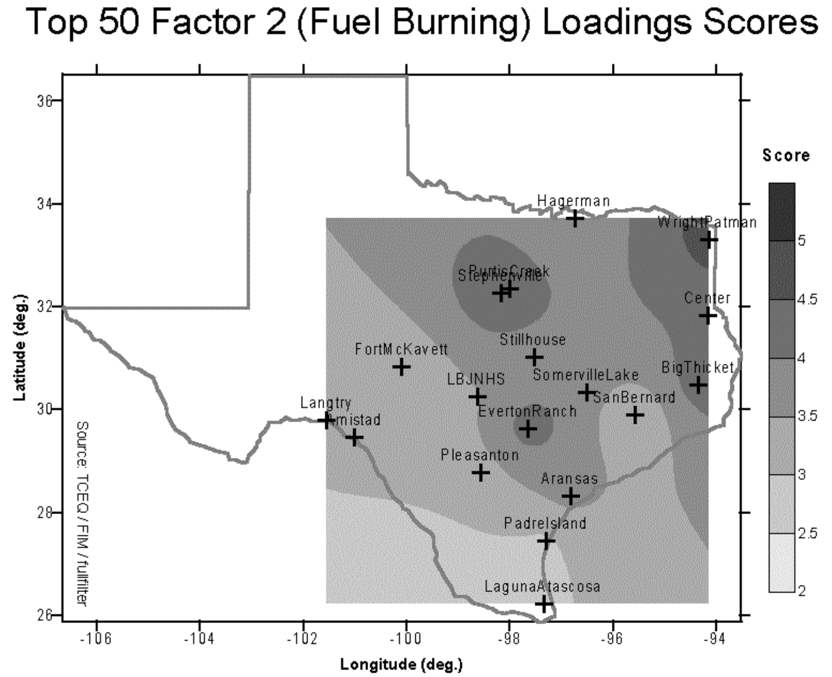


Figure 2 - Demonstrates the location of the top 50 loadings related to coal burning.

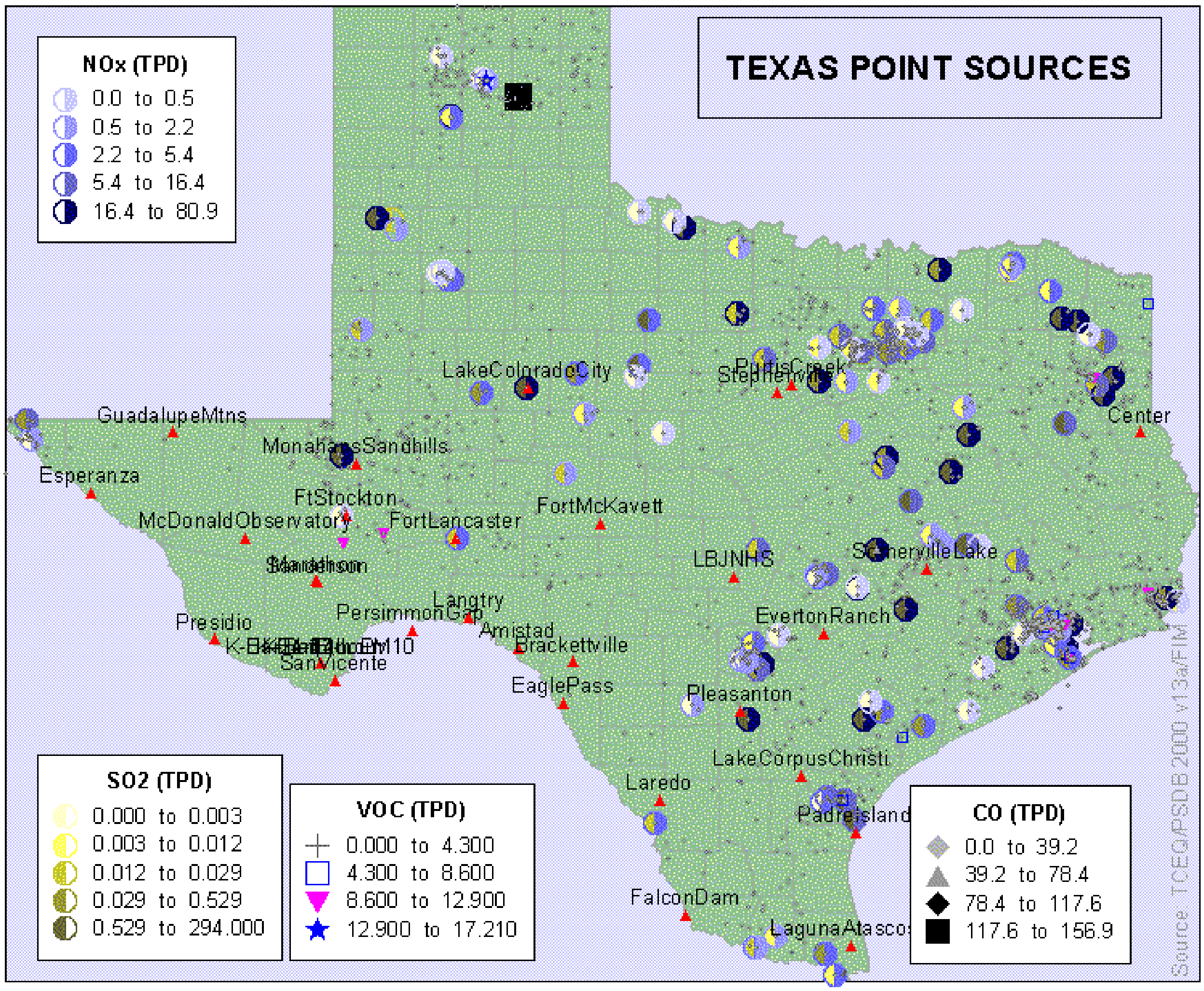


Figure 3 – Shows various types of point sources throughout Texas with respect to sampling site locations. Smaller emissions rates were intentionally depicted small and some sites may not be visible. Each hemisphere depicts either NOx or SO2 concentrations.

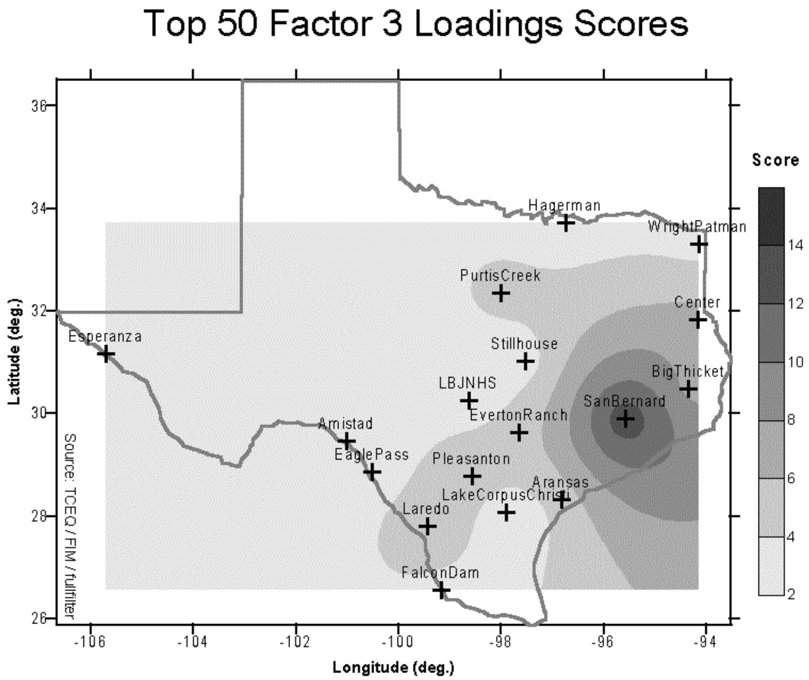


Figure 4 – Unresolved Factor.

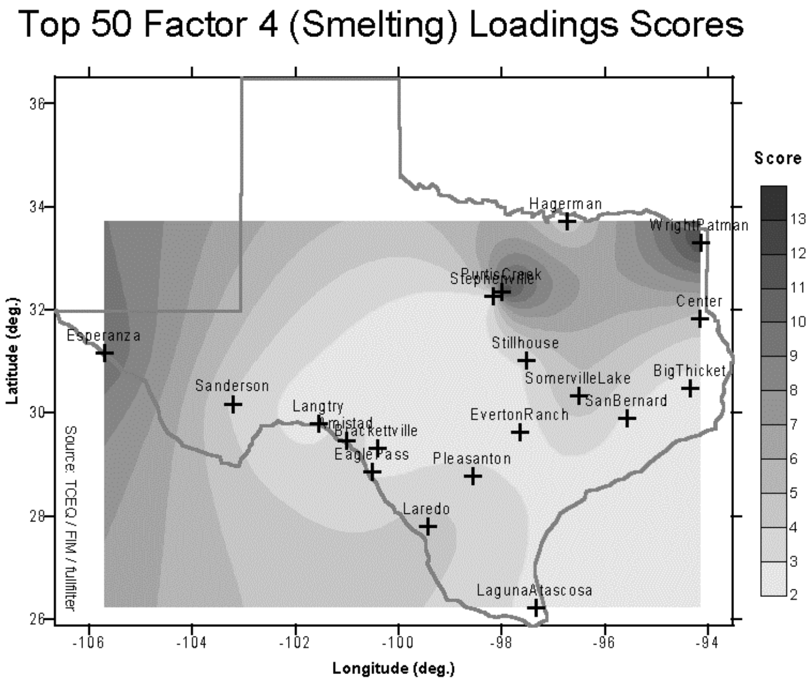


Figure 5 – Top 50 Factor loadings for Factor 4.

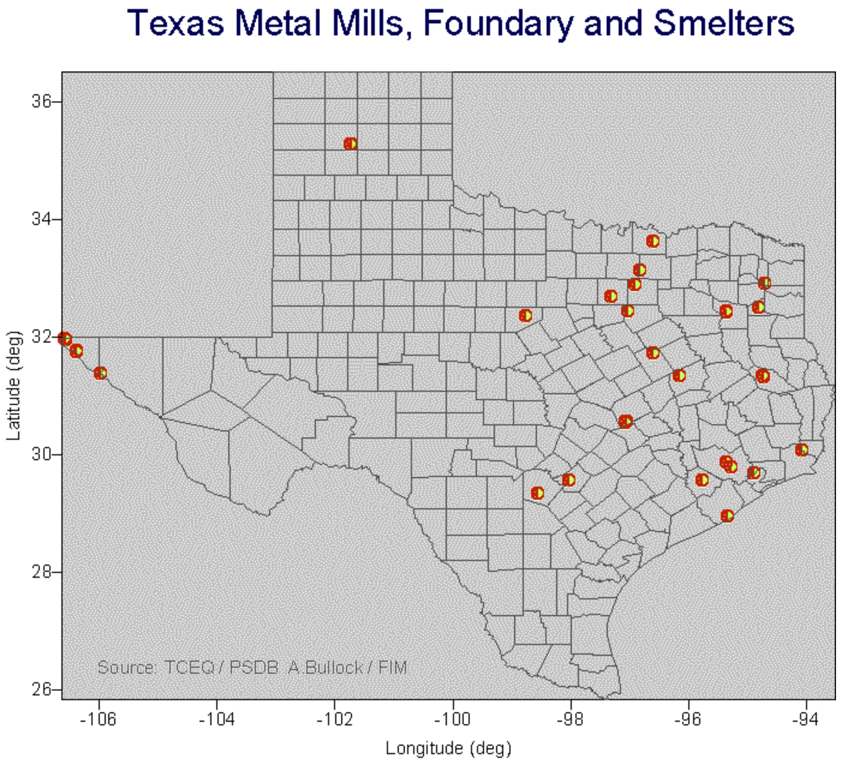

Texas Metal Mills, Foundary and Smelters
36
34
32
30
28
26
Latitude (deg)
-106
-104
-102
-100
-98
-96
-94
Longitude (deg)
Source: TCEQ / PSDB A.Bullock / FIM

Figure 6 - Location of metal mills, foundry and smelters.

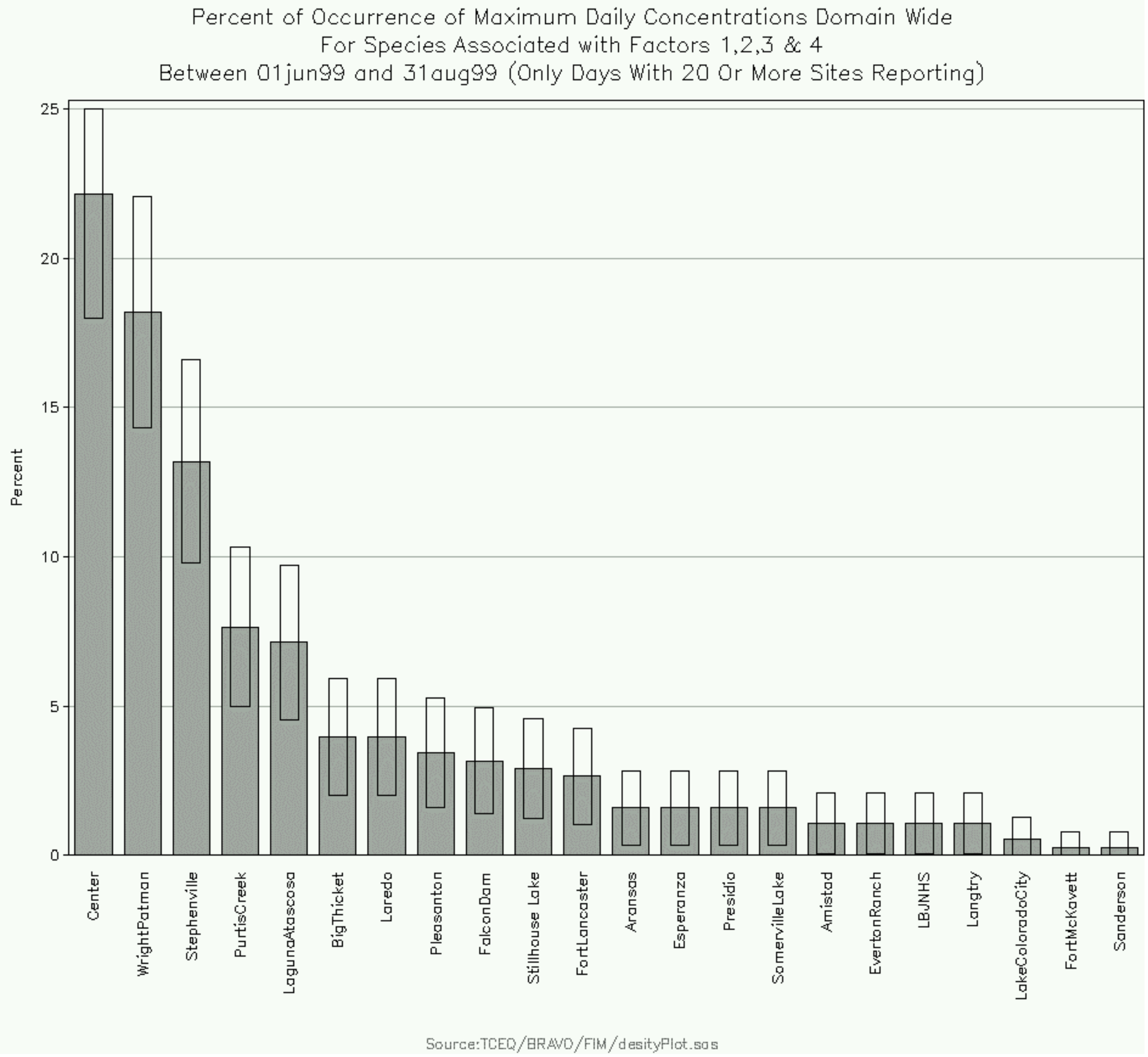


Figure 7 – East Texas sampling sites record more daily maximum concentrations for the early part of the study (01June – 31Aug) than the rest of Texas sampling sites (error bars indicate a 95% C.I.).

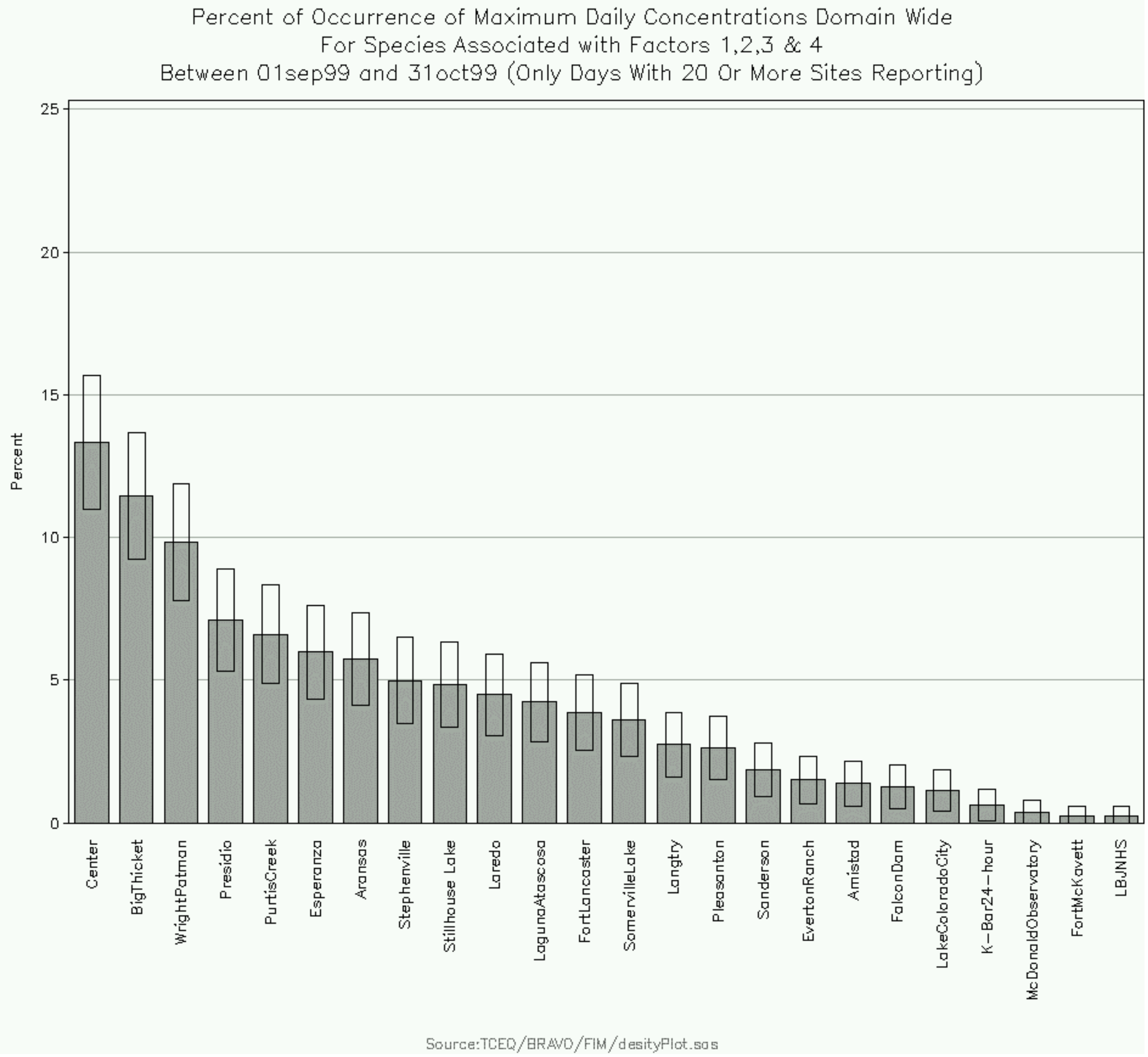


Figure 8 - East Texas sampling sites record more daily maximum concentrations for the late part of the study (September 1 – October 31) than the rest of Texas sampling sites (error bars indicate a 95% C.I.).

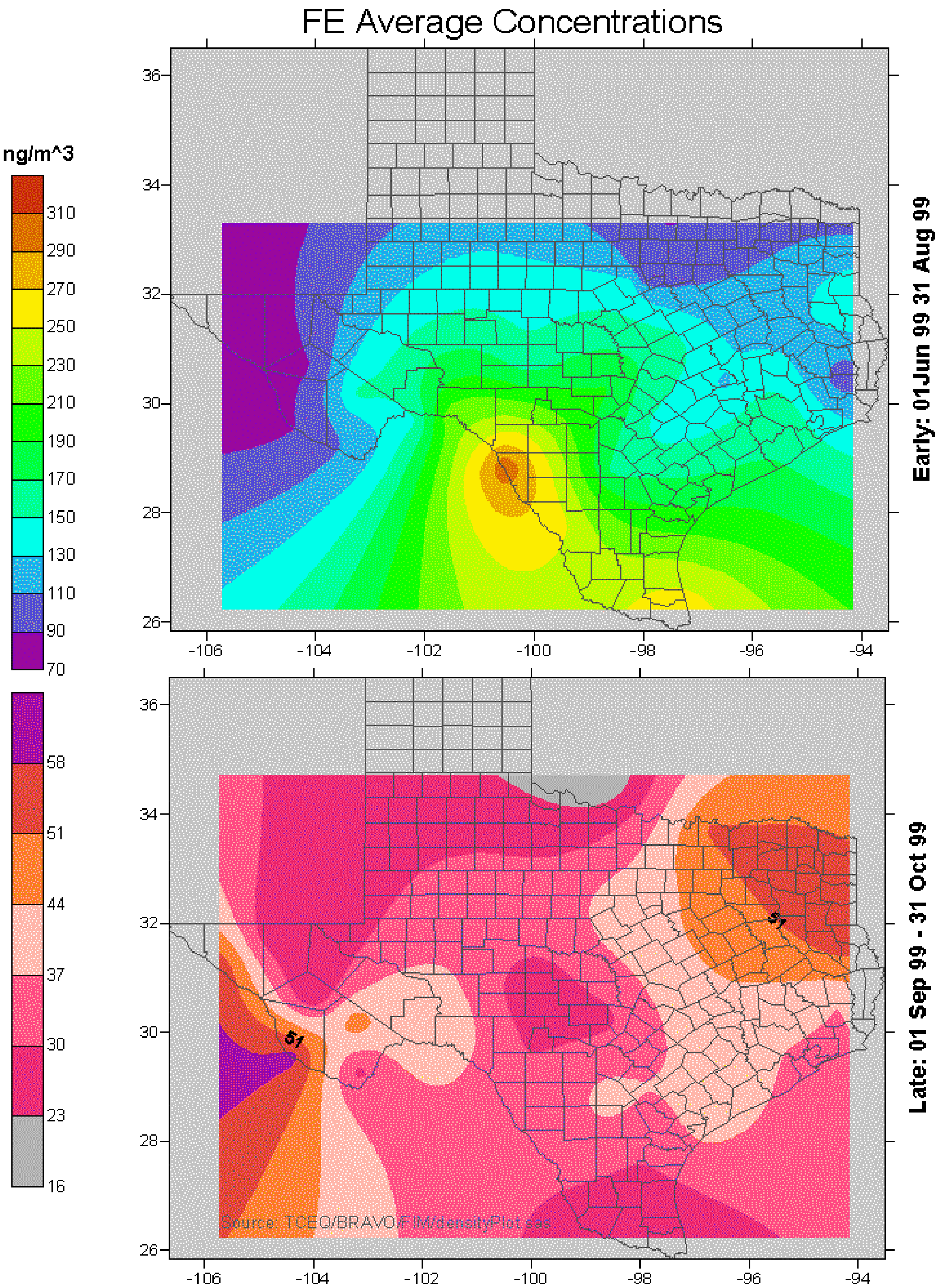

FE Average Concentrations
ng/m^3
310
290
270
250
230
210
190
170
150
130
110
90
70
58
51
44
37
30
23
16
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 31 Oct 99
51
51
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 9 – Illustrates early and late study periods in which the early period's concentrations are higher. Averages contain at least 15 observations.

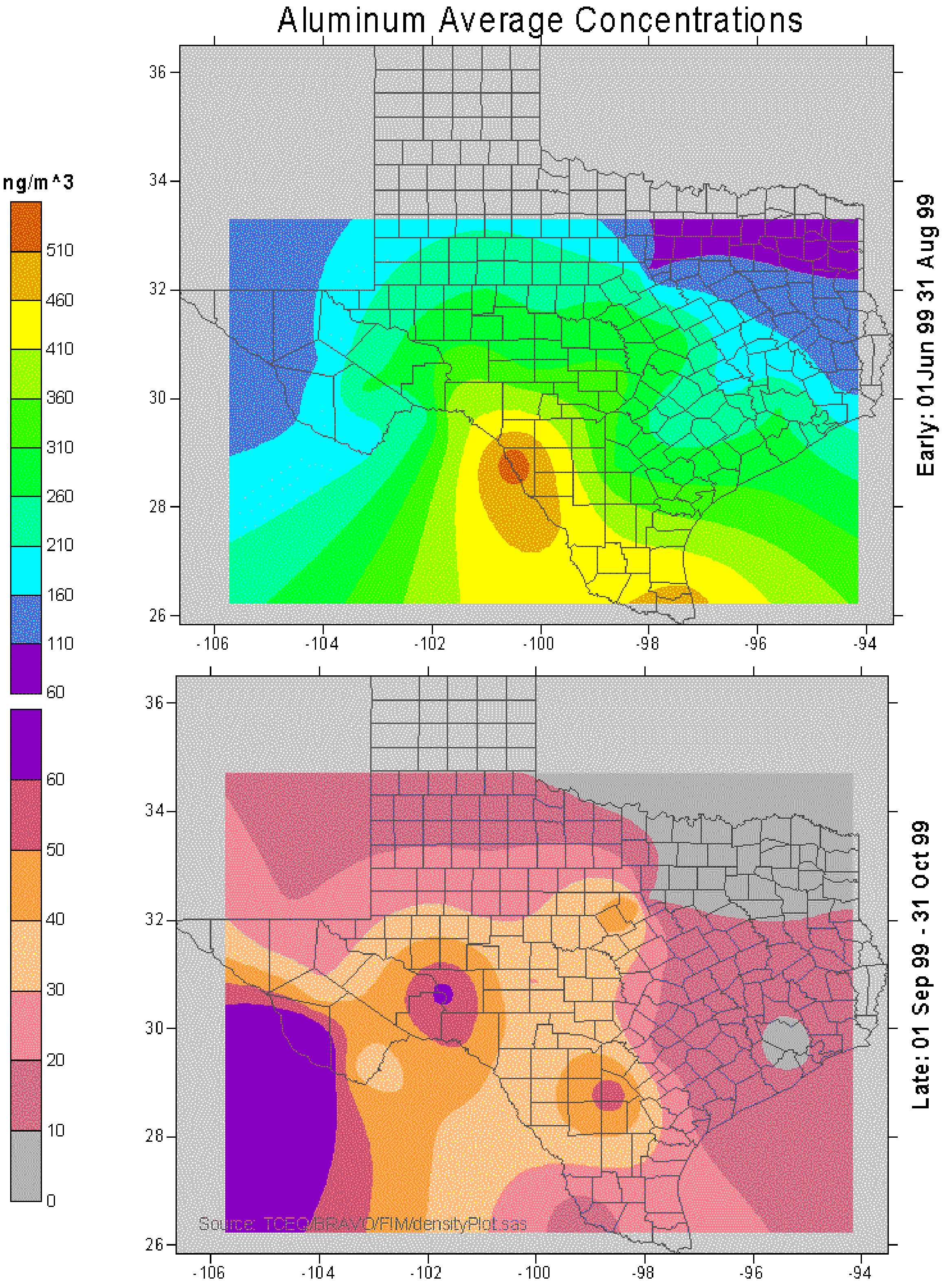

Aluminum Average Concentrations
ng/m^3
510
460
410
360
310
260
210
160
110
60
60
50
40
30
20
10
0
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 31 Oct 99
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 10 - Illustrates early and late study periods in which the early period's concentrations are higher. Averages contain at least 15 observations.

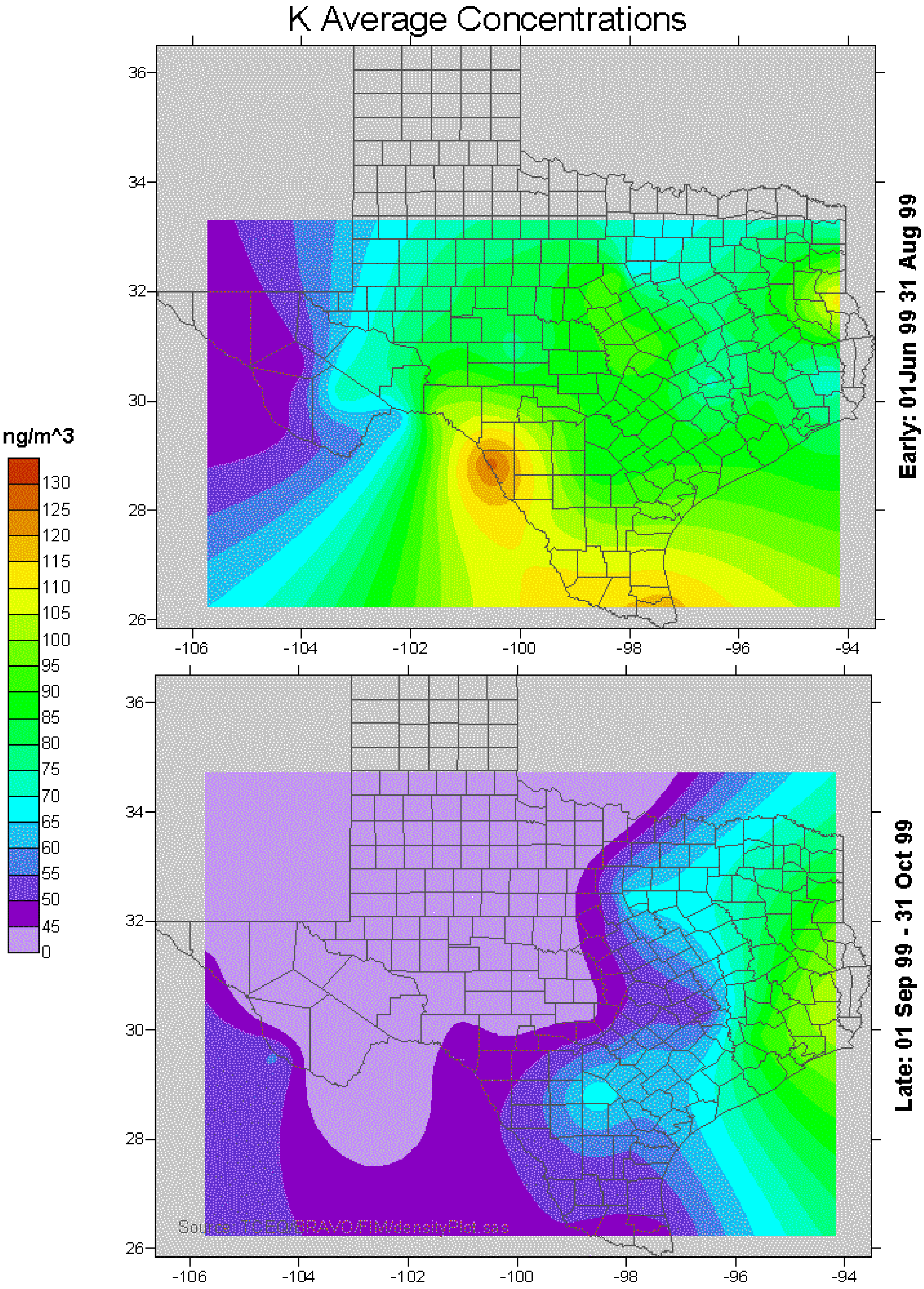


Figure 11 - Averages contain at least 15 observations.

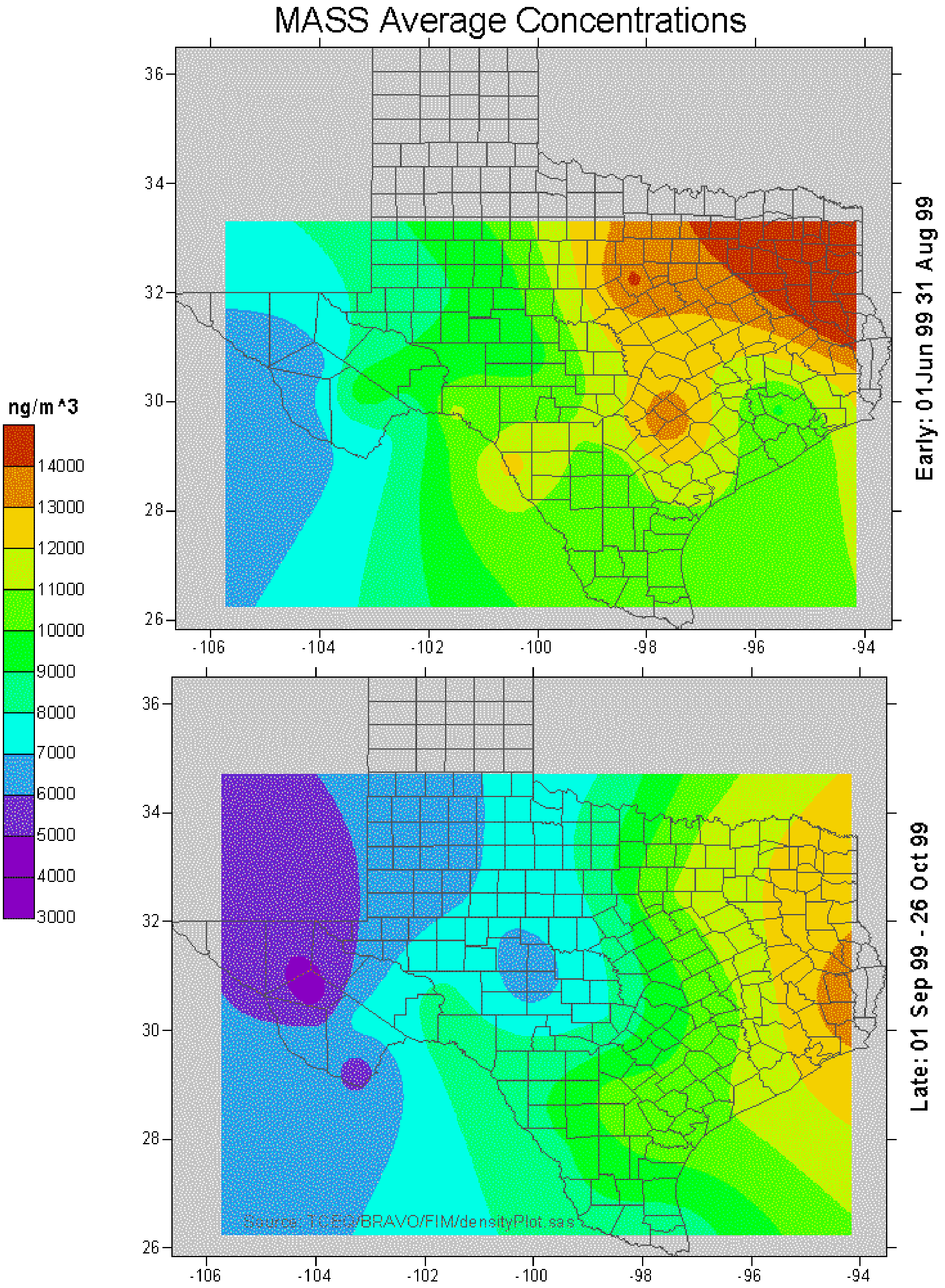

MASS Average Concentrations
ng/m^3
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 26 Oct 99
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 12 – Higher concentrations in east Texas during both study periods. Averages contain at least 15 observations.

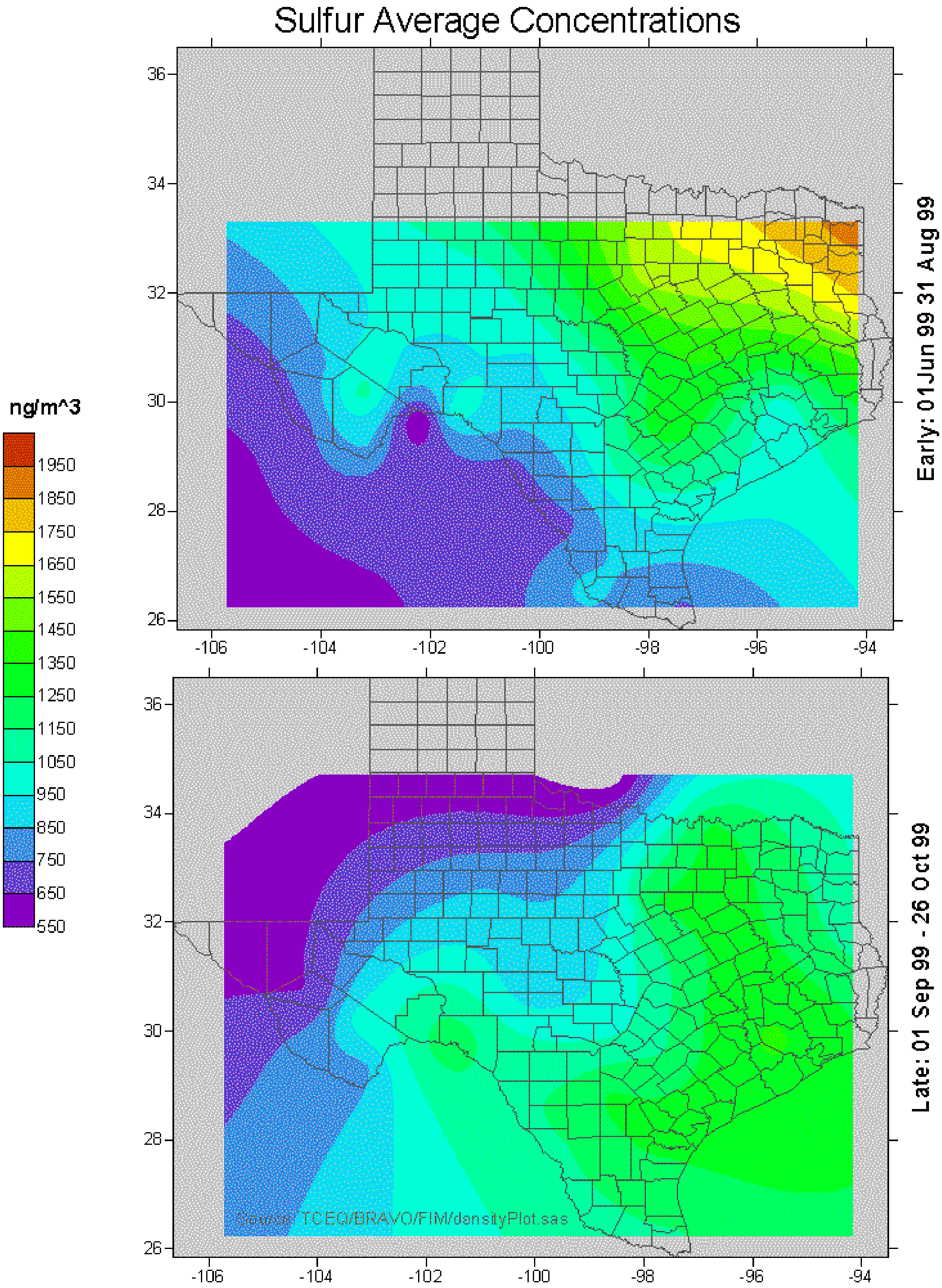

Sulfur Average Concentrations
ng/m^3
1950
1850
1750
1650
1550
1450
1350
1250
1150
1050
950
850
750
650
550
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 26 Oct 99
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 13 – A local maximum at the Langtry/Amistad sites. Averages contain at least 15 observations.

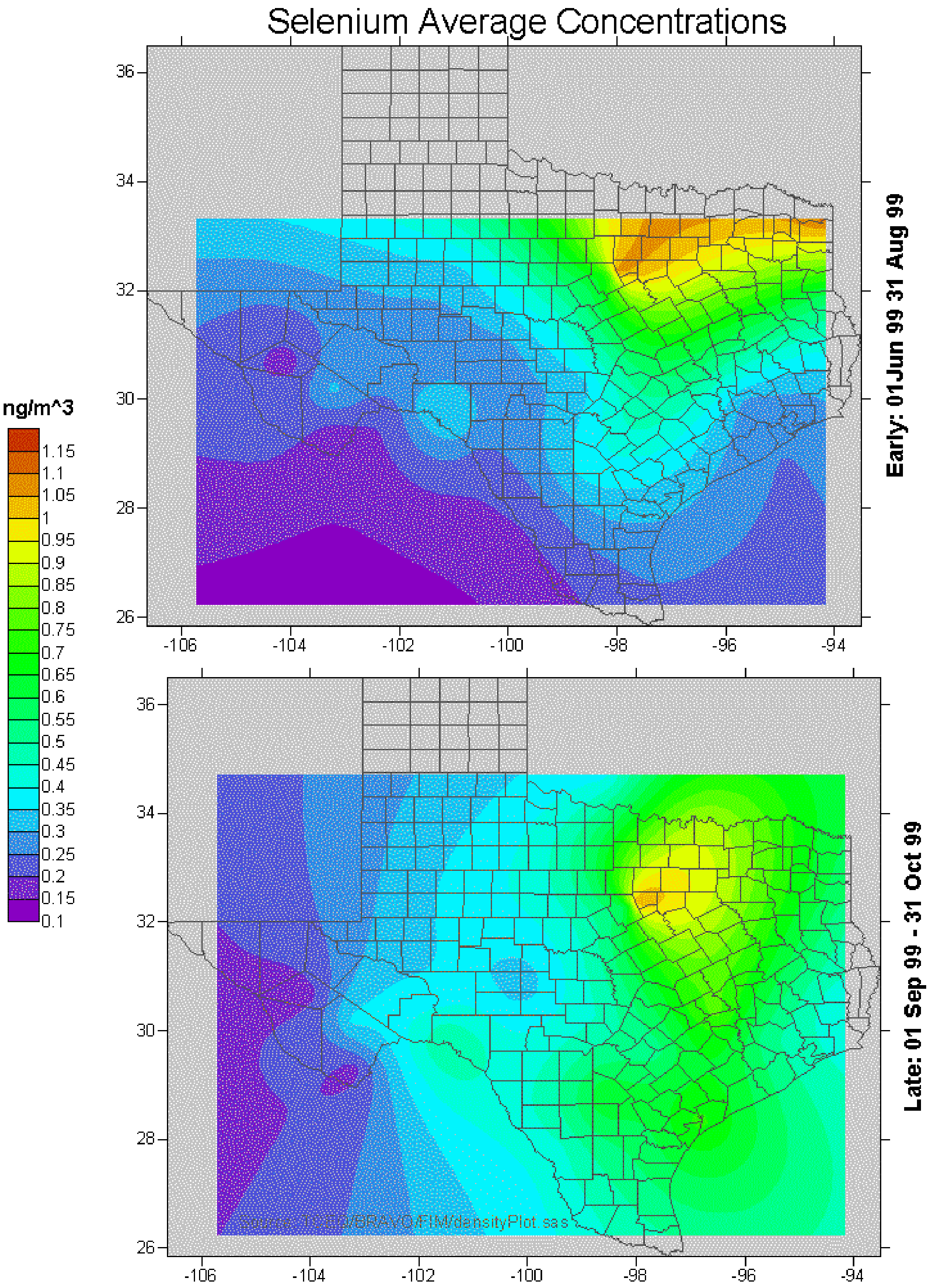

Selenium Average Concentrations
ng/m^3
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 31 Oct 99
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 14 - Averages contain at least 15 observations.

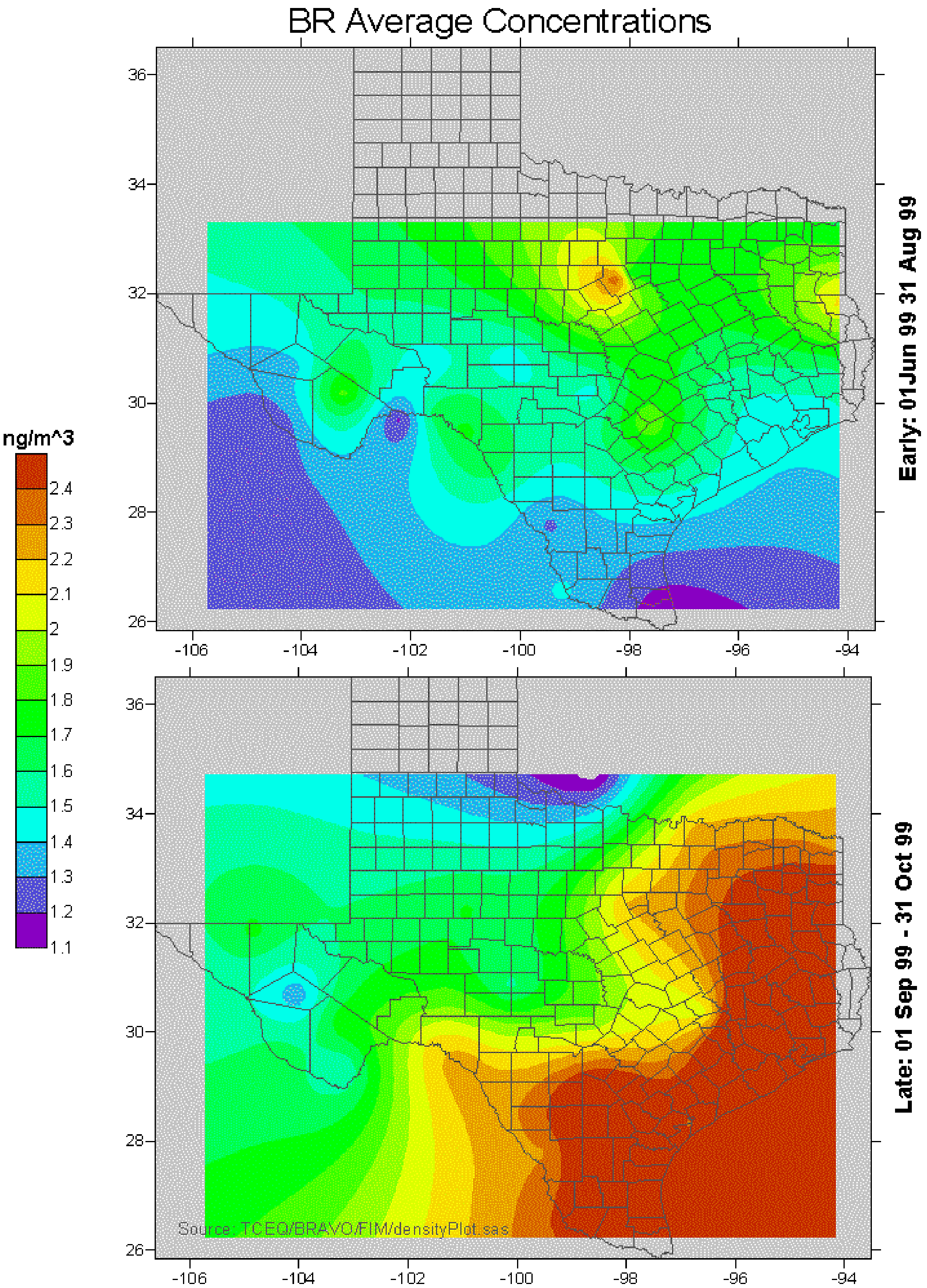
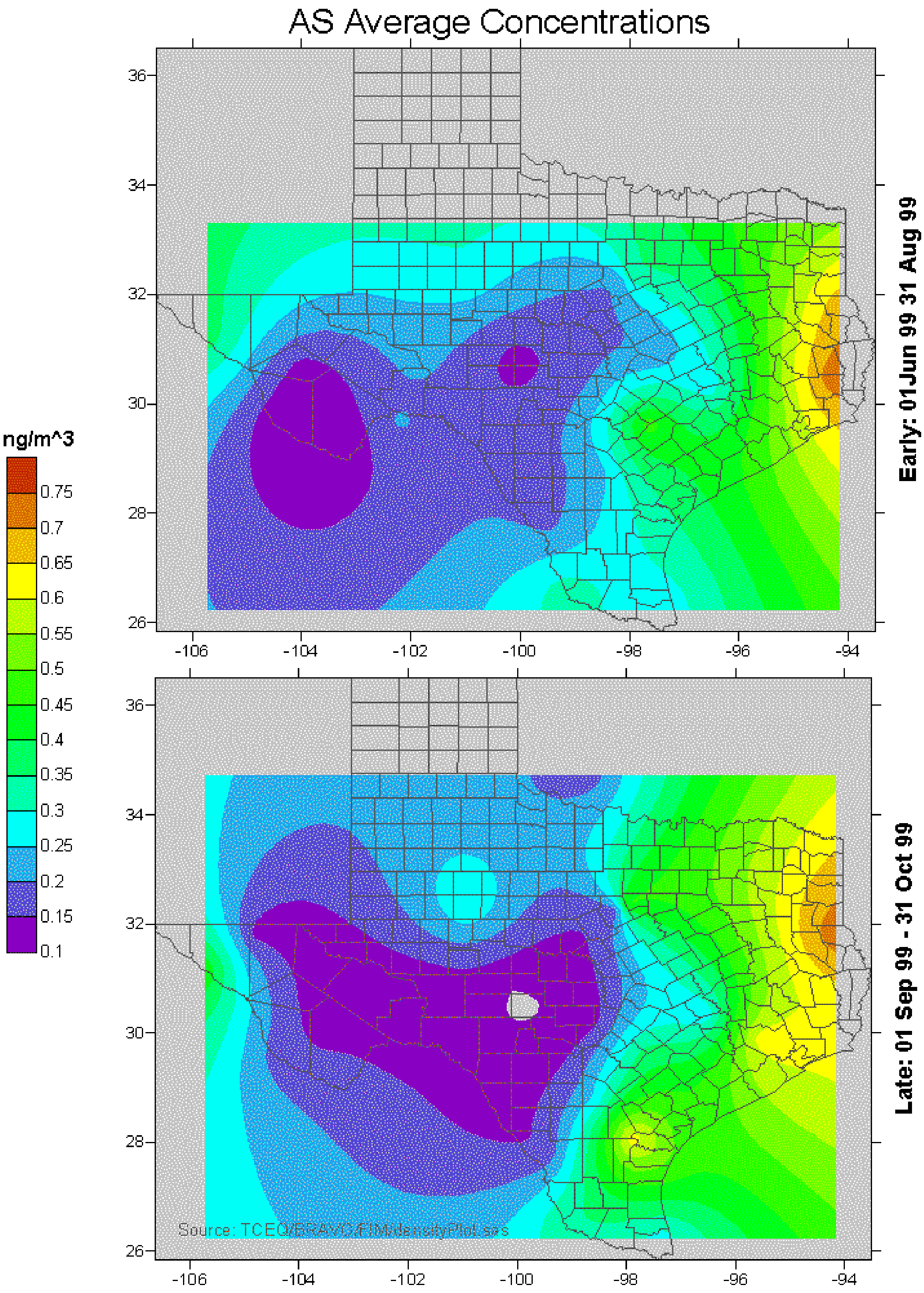

BR Average Concentrations
ng/m^3
2.4
2.3
2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 31 Oct 99
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 15 - Averages contain at least 15 observations.

Source: TCEQ / TAD / FIM. Created on 2/26/04 4:09 PM. Revised on 3/23/04



Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 16 - Averages contain at least 15 observations.

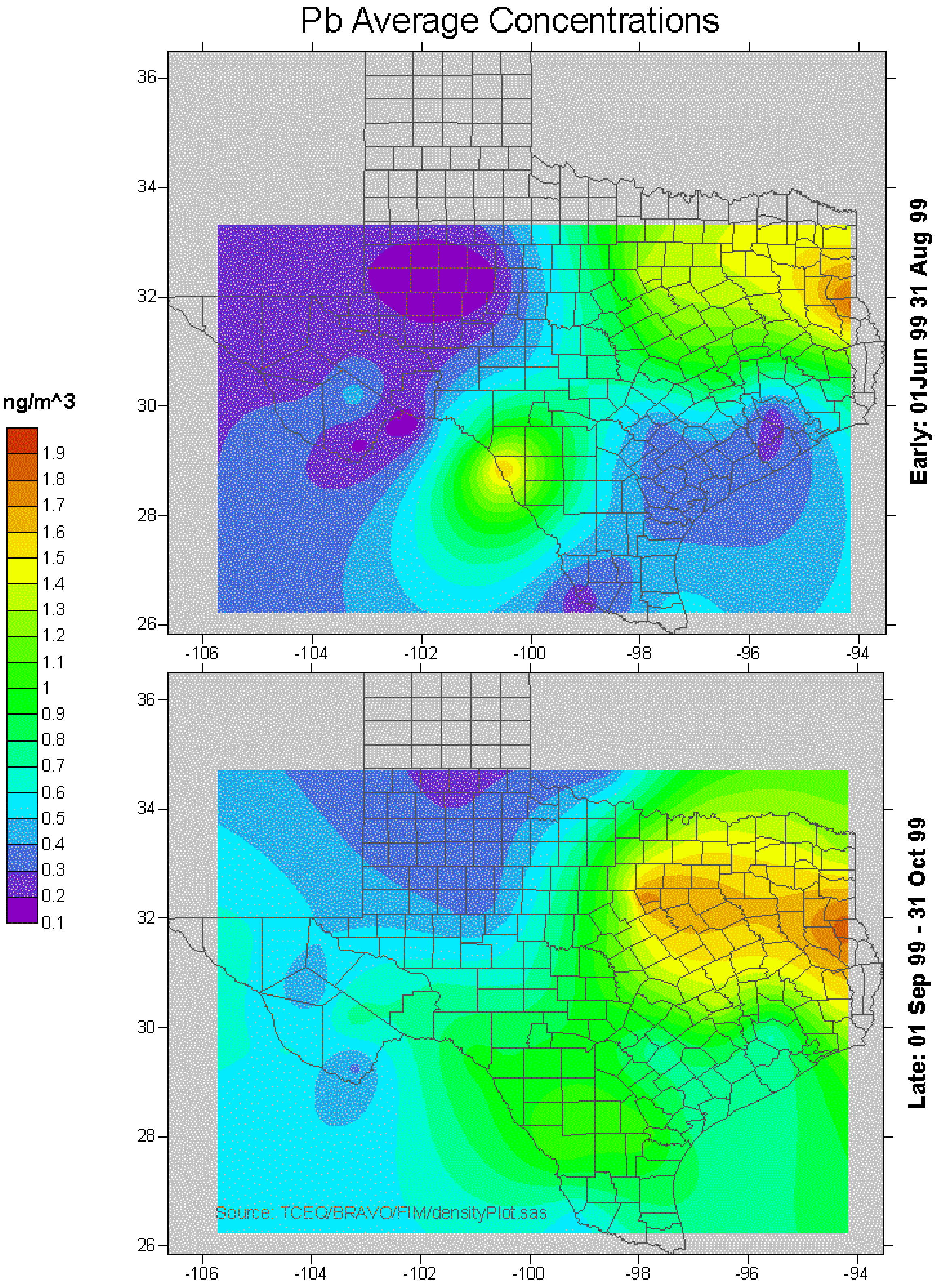
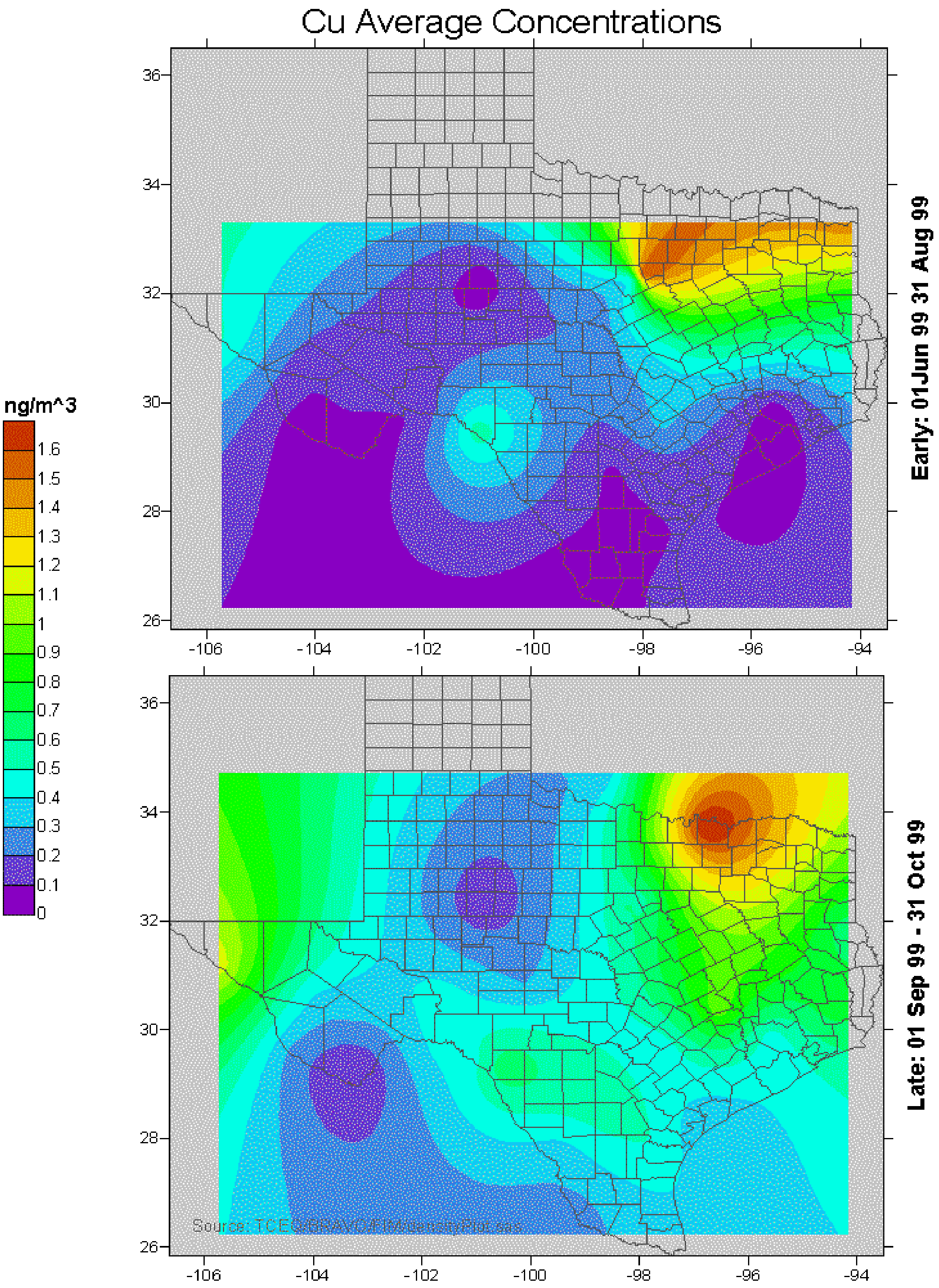

Pb Average Concentrations
ng/m^3
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 31 Oct 99
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 17 - Averages contain at least 15 observations.



Cu Average Concentrations

Early: 01Jun 99 31 Aug 99

Late: 01 Sep 99 - 31 Oct 99

Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 18 - Averages contain at least 15 observations.

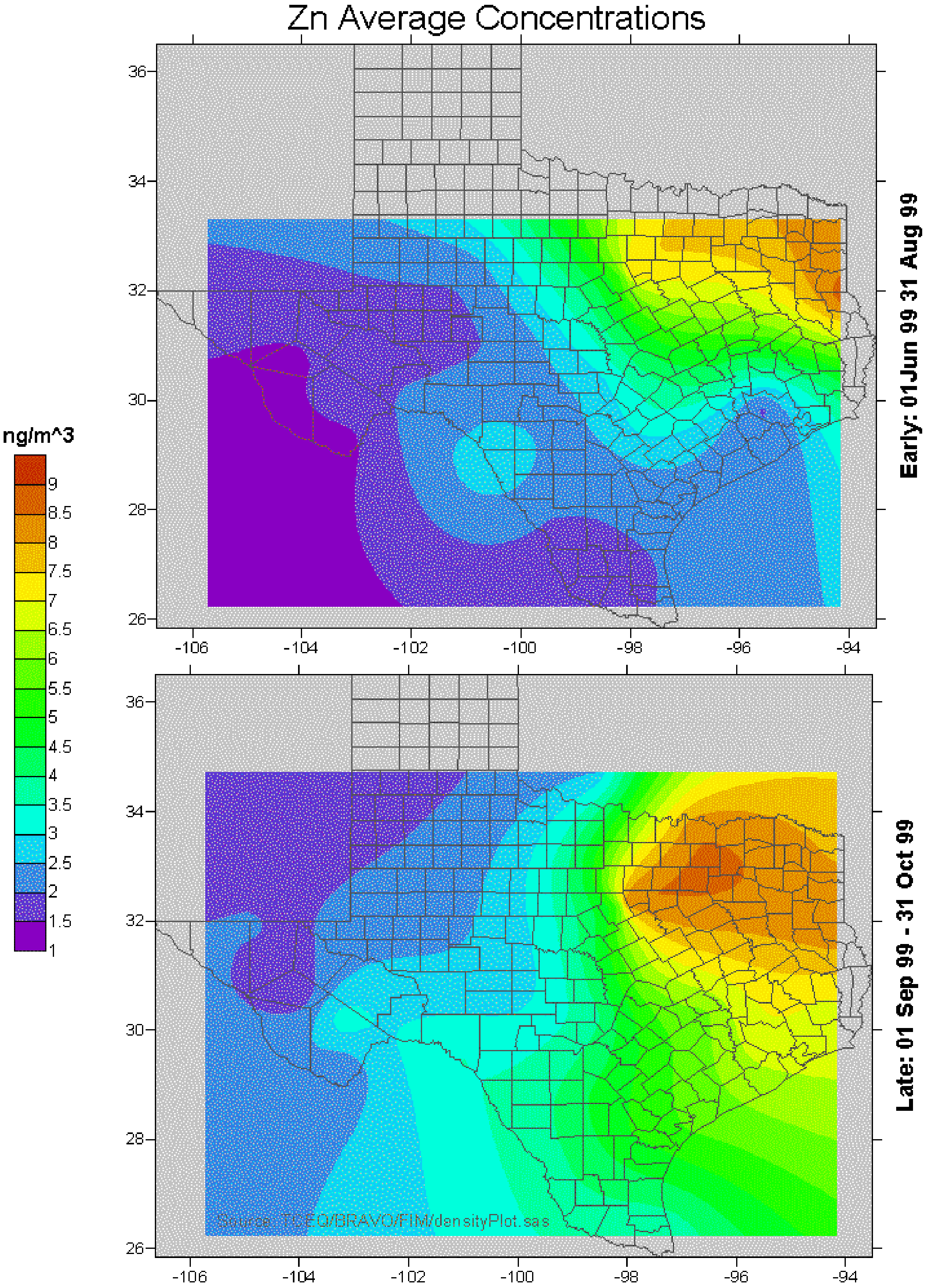

Zn Average Concentrations
ng/m^3
9
8.5
8
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
36
34
32
30
28
26
-106
-104
-102
-100
-98
-96
-94
Early: 01Jun 99 31 Aug 99
Late: 01 Sep 99 - 31 Oct 99
Source: TCEQ/BRAVO/FIM/densityPlot.sas

Figure 19 - Averages contain at least 15 observations.

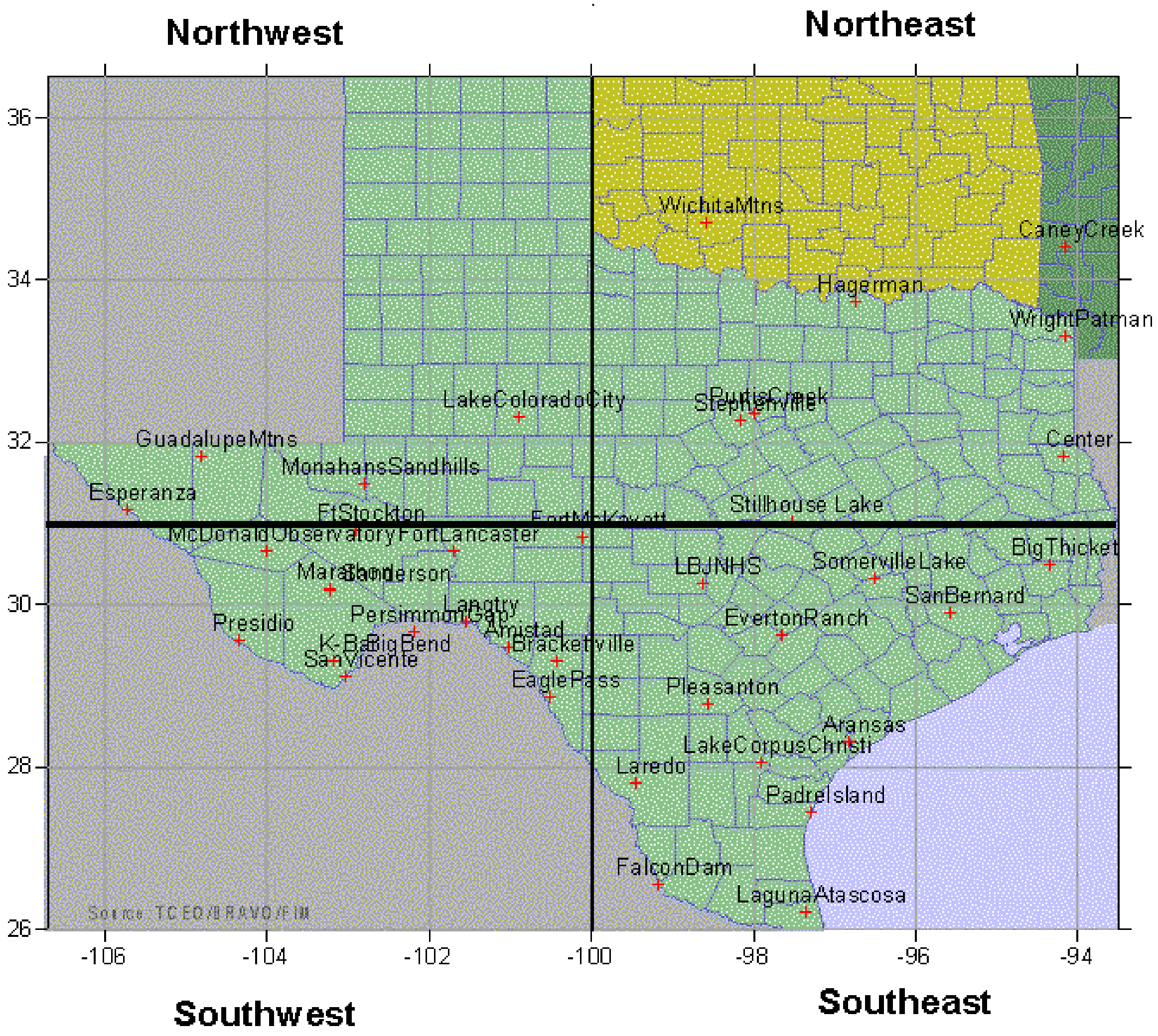


Figure 20 – Illustration of BRAVO sites and regions. Note, Persimmon Gap appears to be located in Mexico but is a U.S. border town. The quadrants regions are arbitrary in size and are used for explaining groups of sites.

XII. References

Gorsuch, Richard L. L. (1974). *Factor Analysis*. Lawrence Erlbaum Associates Inc.

Cattell, R. B. (1958). Extracting the correct number of factors in factor analysis. *Educational and Psychological Measurements*, 18, 791-837.

Cattell, R. B. (1966). The Scree Test for the Number of Factors. *Multivariate Behavior Research*, 1, 245-276.

Rummel, R. J. (1970). *Applied factor analysis*. Evanston, IL: Northwestern University Press.

Texas Commission on Environmental Quality, Point Source Database Version: psdb_all_agg_ems_2000_v13a.sas7bdat.

William Malm, National Park Service, Air Resource Division; personal communication, November 18, 2002.

Georgoulias, Bethany (2002). *Evaluation of the Reconstruction Equation for Predicting Light Extinction at Big Bend during BRAVO,* TCEQ.