COLORADO STATE IMPLEMENTATION PLAN FOR REGIONAL HAZE

TECHNICAL SUPPORT DOCUMENT

Douglas Pass and Ripple Creek Pass Sites Analyses

Colorado Department
of Public Health
and Environment

Prepared by
The Colorado Department of Public Health and Environment
Air Pollution Control Division
4300 Cherry Creek Drive South
Denver, Colorado 80246-1530
(303) 692-3100

June 2007

Douglas Pass and Ripple Creek Pass Sites Analyses

Introduction. The State of Colorado was involved in a project to monitor acid deposition, visibility, and meteorological parameters in northwest Colorado. The project is important to Colorado for defining existing conditions and estimating potential effects on Colorado's Class I areas from current and future air pollution sources. The project had two parts: 1) Shell Exploration and Production Company consulted and collaborated with the State, federal land managers, and the U.S. Geological Survey regarding a monitoring site near the Flat Tops Wilderness Area at Ripple Creek Pass (elevation 9575 ft.). The site began operating approximately December 1, 2002 and is still in operation. The site is funded entirely by Shell. 2) The Colorado Governor's Office of Energy Management and Conservation (OEMC) in collaboration with the Colorado Department of Public Health and Environment/Air Pollution Control Division (CDPHE/APCD) awarded a contract through a competitive bidding process to Air Sciences, Inc. to build and operate a monitoring site approximately 75 miles due west of the Ripple Creek site. The OEMC funded the Douglas Pass site and CDPHE/APCD provided technical support to the OEMC. The site was funded to be in operation approximately 24 months.

The Douglas Pass site (elevation 8200 ft.) consisted of an IMPROVE II particulate monitor and a basic set of meteorological measurements. The site began operating in the fall of 2003 and was shut-down in January 2006. The purpose of this short-term site was to measure regional haze/visibility events as they enter Colorado from the west to better understand regional haze in Colorado's Class I areas. A secondary purpose of the site (related to the first part of the overall project as discussed in the above paragraph) was to provide upwind information in relation to the Ripple Creek site for the period the 2 sites were operating at the same time. This is of interest to the State of Colorado in the context of better understanding air quality at and upwind of the Flat Tops Wilderness Area and potential oil shale development.

The monitoring programs at both Douglas Pass and Ripple Creek Pass consisted of IMPROVE samplers and basic meteorological parameters. Both sites were operated by Air Sciences, Inc. The data are considered to be of similar quality as federally operated IMPROVE sites. The same protocols and equipment were utilized.

Douglas Pass is within 25 miles of the Utah/Colorado border. The site was intentionally selected to be close to the border in order to measure haze as it entered Colorado. The following analyses explore:

- Correlations: relationships between this site and other monitoring sites in Colorado;
- Back Trajectories: where dirty air that was monitored at these sites came from; and
- PMF: possible sources of dirty air at the sites.

The analysis is conducted within a general hypothesis/framework stated below:

Colorado is generally a downwind state in relation to other states in the

West. Haze from upwind states enters Colorado and is a significant contributor to visibility impairment as measured at Colorado monitoring sites. While modeling exercises have been conducted to examine the issue of what source categories, regions and states are contributing to haze at Colorado Class I areas, this project was intended to be monitoring-based. Douglas Pass measurements represent haze entering Colorado and it can be utilized to quantify the influence of out-of-state haze on Colorado Class I areas.

Correlations. Paired Pearson product-moment correlations and significance levels were computed for data collected 2003-2005. Pairs of correlations will have more or less data in common depending on when sites operated during this period, missing data, operational problems, etc.

Aerosol Extinction, Worst Days. The worst 20% aerosol extinction days were determined for Ripple Creek Pass. A squared Pearson correlation between Ripple Creek Pass and Douglas Pass will represent the amount of variation at Ripple Creek Pass (downwind) that can be associated with variation in measurements at Douglas Pass (upwind). Table I indicates a high correlation of .969 between the two sites. This is an r^2 of .94 or 94%.

Table 1
Worst Day Aerosol Extinction 2003-2005

		Douglas Pass	Ripple Creek Pass
Douglas Pass	Pearson Correlation	1	.969(**)
	Sig. (2-tailed)		.000
	N	47	47
Ripple Creek Pass	Pearson Correlation	.969(**)	1
	Sig. (2-tailed)	.000	
	N	47	47

** Correlation is significant at the 0.01 level (2-tailed).

Extinction, All Days. Table 2 presents an entire matrix of aerosol extinction correlations between all sites that had data during some part of 2003-2005. Of particular interest is the Douglas Pass column (highlighted in yellow). One can look down it to see correlations with it ranging from .961 (Ripple Creek) to -.236 (Shamrock Mines). Generally, sites closer to and with a similar elevation of Douglas Pass have the higher correlations. Canyonlands NP in Utah was included to examine how related a site that is further to the west (e.g., upwind) compares to Colorado sites.

Tables 3 thru 8 present correlations between Douglas Pass and the other sites for each of the important visibility impairing species.

Table 2
Aerosol Extinction Correlations/All Valid Data 2003-2005

		Douglas Pass	Ripple Creek Pass	Canyonlands NP	Great Sand Dunes NP	Mesa Verde NP	Mt. Zirkel WA	Rocky Mountain NP	Shamrock Mines	Weminuche WA	White River
Douglas Pass	Pearson Correlation	1	.961(**)	.382(**)	.051	.813(**)	.799(**)	.211(*)	-.236	.859(**)	.543(**)
	Sig. (2-tailed)		.000	.000	.591	.000	.000	.023	.143	.000	.000
	N	220	191	108	114	116	104	116	40	112	111
Ripple Creek Pass	Pearson Correlation	.961(**)	1	.565(**)	.516(**)	.751(**)	.912(**)	.389(**)	.409(**)	.847(**)	.765(**)
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.003	.000	.000
	N	191	286	158	188	164	167	174	51	171	186
Canyonlands NP	Pearson Correlation	.382(**)	.565(**)	1	.408(**)	.479(**)	.430(**)	.348(**)	.220	.365(**)	.465(**)
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.114	.000	.000
	N	108	158	225	217	209	199	217	53	214	220
Great Sand Dunes NP	Pearson Correlation	.051	.516(**)	.408(**)	1	.623(**)	.539(**)	.545(**)	.659(**)	.547(**)	.605(**)
	Sig. (2-tailed)	.591	.000	.000		.000	.000	.000	.000	.000	.000
	N	114	188	217	251	221	221	231	61	226	245
Mesa Verde NP	Pearson Correlation	.813(**)	.751(**)	.479(**)	.623(**)	1	.714(**)	.422(**)	.697(**)	.740(**)	.640(**)
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000
	N	116	164	209	221	230	205	223	53	218	223
Mt. Zirkel WA	Pearson Correlation	.799(**)	.912(**)	.430(**)	.539(**)	.714(**)	1	.594(**)	.559(**)	.731(**)	.790(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000
	N	104	167	199	221	205	231	215	54	207	224
Rocky Mountain NP	Pearson Correlation	.211(*)	.389(**)	.348(**)	.545(**)	.422(**)	.594(**)	1	.440(**)	.356(**)	.740(**)
	Sig. (2-tailed)	.023	.000	.000	.000	.000	.000		.001	.000	.000
	N	116	174	217	231	223	215	240	55	228	234
Shamrock Mines	Pearson Correlation	-.236	.409(**)	.220	.659(**)	.697(**)	.559(**)	.440(**)	1	.131	.530(**)
	Sig. (2-tailed)	.143	.003	.114	.000	.000	.000	.001		.356	.000
	N	40	51	53	61	53	54	55	67	52	65
Weminuche WA	Pearson Correlation	.859(**)	.847(**)	.365(**)	.547(**)	.740(**)	.731(**)	.356(**)	.131	1	.710(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.356		.000
	N	112	171	214	226	218	207	228	52	235	229
White River	Pearson Correlation	.543(**)	.765(**)	.465(**)	.605(**)	.640(**)	.790(**)	.740(**)	.530(**)	.710(**)	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	N	111	186	220	245	223	224	234	65	229	254

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Table 3
Sulfur Correlations/All Valid Data 2003-2005

		Douglas Pass	Ripple Creek Pass	Canyonlands NP	Great Sand Dunes NP	Mesa Verde NP	Mt. Zirkel WA	Rocky Mountain NP	Shamrock Mines	Weminuche WA	White River
Douglas Pass	Pearson Correlation	1	.804(**)	.757(**)	.534(**)	.580(**)	.748(**)	.594(**)	.596(**)	.635(**)	.690(**)
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	263	237	230	224	223	212	227	143	229	229
Ripple Creek Pass	Pearson Correlation	.804(**)	1	.689(**)	.618(**)	.650(**)	.824(**)	.695(**)	.629(**)	.683(**)	.802(**)
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	237	383	288	288	275	271	288	140	290	293
Canyonlands NP	Pearson Correlation	.757(**)	.689(**)	1	.630(**)	.747(**)	.658(**)	.541(**)	.708(**)	.733(**)	.713(**)
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000	.000	.000
	N	230	288	334	323	312	306	322	145	325	329
Great Sand Dunes NP	Pearson Correlation	.534(**)	.618(**)	.630(**)	1	.769(**)	.635(**)	.624(**)	.831(**)	.812(**)	.744(**)
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000
	N	224	288	323	333	309	306	321	135	324	328
Mesa Verde NP	Pearson Correlation	.580(**)	.650(**)	.747(**)	.769(**)	1	.602(**)	.491(**)	.907(**)	.884(**)	.761(**)
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000
	N	223	275	312	309	321	297	309	136	313	315
Mt. Zirkel WA	Pearson Correlation	.748(**)	.824(**)	.658(**)	.635(**)	.602(**)	1	.758(**)	.628(**)	.667(**)	.814(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000
	N	212	271	306	306	297	317	305	129	310	311
Rocky Mountain NP	Pearson Correlation	.594(**)	.695(**)	.541(**)	.624(**)	.491(**)	.758(**)	1	.494(**)	.548(**)	.711(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.000
	N	227	288	322	321	309	305	333	139	324	327
Shamrock Mines	Pearson Correlation	.596(**)	.629(**)	.708(**)	.831(**)	.907(**)	.628(**)	.494(**)	1	.949(**)	.791(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000		.000	.000
	N	143	140	145	135	136	129	139	146	141	144
Weminuche WA	Pearson Correlation	.635(**)	.683(**)	.733(**)	.812(**)	.884(**)	.667(**)	.548(**)	.949(**)	1	.814(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000		.000
	N	229	290	325	324	313	310	324	141	336	330
White River	Pearson Correlation	.690(**)	.802(**)	.713(**)	.744(**)	.761(**)	.814(**)	.711(**)	.791(**)	.814(**)	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	N	229	293	329	328	315	311	327	144	330	339

*Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

.

Sulfur, All Days. Table 3 contains correlations coefficients for sulfur between pairs of sites. With particular attention to the Douglas Pass column, there are no low or negative correlations between Douglas Pass and any other site. Generally 25-50% of sulfur measured at Douglas Pass appears to be associated with sulfur measured at other sites.

Nitrate, All Days. Table 4 has a strong relationship between Douglas Pass and Ripple Creek Pass nitrate but very low, no, or negative correlation coefficients with other sites.

Elemental Carbon, All Days. Table 5's highest correlation is with the White River site (r=0.78), not with Ripple Creek (0.58). Other correlations range from negative (Shamrock Mines), to negligible (Rocky Mountain), to moderate (remaining sites).

Organic Carbon, All Days. Table 6 shows very high correlations with Douglas Pass, averaging around 0.90, for all sites except Shamrock Mines (0.20).

Fine Mass, All Days. Table 7's correlation pairs with Douglas Pass are mostly very low (Shamrock Mines, White River, Rocky Mountain NP, Mt. Zirkel WA, Mesa Verde NP, Great Sand Dunes). Two sites have a moderate relationship with Douglas Pass fine mass (Weminuche WA, Canyonlands NP) and only one site, Ripple Creek, has a high correlation. Examining the rest of the correlation matrix in Table 7 indicates that Colorado sites are fairly highly inter-correlated.

Coarse Mass, All Days. Overall, the pairs in Table 8 are not very correlated. The highest is between Douglas Pass and Ripple Creek (0.88). In contrast with the fine mass correlation matrix, few sites have much relation with each other suggesting very local source areas for coarse mass.

Summary of Correlations of other IMPROVE monitors with Douglas Pass. These monitoring-based results are consistent with other evidence. Using Douglas Pass as an indicator/predictor of regional haze originating out-of-state, shows:

- Worst days aerosol extinction at Ripple Creek Pass is highly associated (r^2=94%) with Douglas Pass. That is, the worst haze at Ripple Creek is "explained" by haze coming from out-of-state.
- All days aerosol extinction at Douglas Pass, and out-of-state haze, is highly associated with the sites closest to it and poorly correlated to sites at greater distance.
- Douglas Pass sulfur is consistently well correlated to all sites with coefficients ranging from 0.53 to 0.80. This might indicate that sulfur variability in Colorado is quite regional in scale.
- In contrast, Douglas Pass nitrate is poorly correlated at all sites except Ripple Creek indicating the importance of more local source areas.
- Elemental carbon measurements at Douglas Pass have low correlations with all sites except White River (0.78).
- Organic carbon at Douglas Pass is very high correlated with other sites. Compared to the other visibility impairing species, OC displays the highest set of correlations. This underlines the strong regional and episodic nature of organic carbon variability in Colorado.
- Fine mass at Douglas Pass has a strong relationship with Ripple Creek Pass and a moderate correlation with Weminunche WA but only small or no association with fine mass at other sites in Colorado. Fine mass at other sites in Colorado does seem to be varying over time in a fairly correlated manner.
- Coarse Mass at Douglas Pass is again correlated well with Ripple Creek but shows little to no relation to other Colorado sites.

Table 4
Nitrate Correlations/All Valid Data 2003-2005

		Douglas Pass	Ripple Creek Pass	Canyonlands NP	Great Sand Dunes NP	Mesa Verde NP	Mt. Zirkel WA	Rocky Mountain NP	Shamrock Mines	Weminuche WA	White River
Douglas Pass	Pearson Correlation	1	.873(**)	-.020	-.323(**)	.073	.094	.052	-.059	.133(*)	.109
	Sig. (2-tailed)		.000	.764	.000	.281	.175	.438	.485	.044	.101
	N	239	221	228	228	223	210	225	142	229	226
Ripple Creek Pass	Pearson Correlation	.873(**)	1	.261(**)	.597(**)	.550(**)	.806(**)	.184(**)	.132	.696(**)	.831(**)
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.002	.112	.000	.000
	N	221	299	275	289	271	262	273	146	280	288
Canyonlands NP	Pearson Correlation	-.020	.261(**)	1	.434(**)	.360(**)	.198(**)	.041	.288(**)	.283(**)	.226(**)
	Sig. (2-tailed)	.764	.000		.000	.000	.001	.462	.000	.000	.000
	N	228	275	335	329	319	299	321	149	329	327
Great Sand Dunes NP	Pearson Correlation	-.323(**)	.597(**)	.434(**)	1	.416(**)	.486(**)	.202(**)	.517(**)	.452(**)	.569(**)
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000
	N	228	289	329	355	326	314	331	153	339	346
Mesa Verde NP	Pearson Correlation	.073	.550(**)	.360(**)	.416(**)	1	.340(**)	.102	.752(**)	.610(**)	.511(**)
	Sig. (2-tailed)	.281	.000	.000	.000		.000	.068	.000	.000	.000
	N	223	271	319	326	332	297	319	141	326	324
Mt. Zirkel WA	Pearson Correlation	.094	.806(**)	.198(**)	.486(**)	.340(**)	1	.220(**)	.168(*)	.488(**)	.731(**)
	Sig. (2-tailed)	.175	.000	.001	.000	.000		.000	.049	.000	.000
	N	210	262	299	314	297	321	302	138	307	313
Rocky Mountain NP	Pearson Correlation	.052	.184(**)	.041	.202(**)	.102	.220(**)	1	.143	.100	.183(**)
	Sig. (2-tailed)	.438	.002	.462	.000	.068	.000		.088	.070	.001
	N	225	273	321	331	319	302	336	143	331	329
Shamrock Mines	Pearson Correlation	-.059	.132	.288(**)	.517(**)	.752(**)	.168(*)	.143	1	.258(**)	.309(**)
	Sig. (2-tailed)	.485	.112	.000	.000	.000	.049	.088		.002	.000
	N	142	146	149	153	141	138	143	160	146	156
Weminuche WA	Pearson Correlation	.133(*)	.696(**)	.283(**)	.452(**)	.610(**)	.488(**)	.100	.258(**)	1	.760(**)
	Sig. (2-tailed)	.044	.000	.000	.000	.000	.000	.070	.002		.000
	N	229	280	329	339	326	307	331	146	344	337
White River	Pearson Correlation	.109	.831(**)	.226(**)	.569(**)	.511(**)	.731(**)	.183(**)	.309(**)	.760(**)	1
	Sig. (2-tailed)	.101	.000	.000	.000	.000	.000	.001	.000	.000	
	N	226	288	327	346	324	313	329	156	337	354

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Table 5
Elemental Carbon Correlations/All Valid Data 2003-2005

		Douglas Pass	Ripple Creek Pass	Canyonlands NP	Great Sand Dunes NP	Mesa Verde NP	Mt. Zirkel WA	Rocky Mountain NP	Shamrock Mines	Weminuche WA	White River
Douglas Pass	Pearson Correlation	1	.582(**)	.408(**)	.313(**)	.344(**)	.399(**)	.115	-.092	.539(**)	.777(**)
	Sig. (2-tailed)		.000	.000	.001	.000	.000	.205	.551	.000	.000
	N	225	202	122	120	121	111	122	44	120	117
Ripple Creek Pass	Pearson Correlation	.582(**)	1	.662(**)	.589(**)	.577(**)	.782(**)	.513(**)	.090	.499(**)	.661(**)
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.526	.000	.000
	N	202	299	181	201	178	182	184	52	185	196
Canyonlands NP	Pearson Correlation	.408(**)	.662(**)	1	.514(**)	.556(**)	.771(**)	.260(**)	.077	.502(**)	.651(**)
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.577	.000	.000
	N	122	181	241	238	233	221	236	55	236	235
Great Sand Dunes NP	Pearson Correlation	.313(**)	.589(**)	.514(**)	1	.495(**)	.622(**)	.302(**)	.118	.499(**)	.496(**)
	Sig. (2-tailed)	.001	.000	.000		.000	.000	.000	.352	.000	.000
	N	120	201	238	258	235	238	241	64	241	252
Mesa Verde NP	Pearson Correlation	.344(**)	.577(**)	.556(**)	.495(**)	1	.610(**)	.297(**)	.059	.536(**)	.538(**)
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.676	.000	.000
	N	121	178	233	235	238	222	233	53	233	232
Mt. Zirkel WA	Pearson Correlation	.399(**)	.782(**)	.771(**)	.622(**)	.610(**)	1	.429(**)	.217	.481(**)	.776(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.105	.000	.000
	N	111	182	221	238	222	241	224	57	225	235
Rocky Mountain NP	Pearson Correlation	.115	.513(**)	.260(**)	.302(**)	.297(**)	.429(**)	1	.128	.243(**)	.242(**)
	Sig. (2-tailed)	.205	.000	.000	.000	.000	.000		.351	.000	.000
	N	122	184	236	241	233	224	243	55	239	238
Shamrock Mines	Pearson Correlation	-.092	.090	.077	.118	.059	.217	.128	1	-.085	.115
	Sig. (2-tailed)	.551	.526	.577	.352	.676	.105	.351		.550	.360
	N	44	52	55	64	53	57	55	67	52	65
Weminuche WA	Pearson Correlation	.539(**)	.499(**)	.502(**)	.499(**)	.536(**)	.481(**)	.243(**)	-.085	1	.600(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.550		.000
	N	120	185	236	241	233	225	239	52	243	238
White River	Pearson Correlation	.777(**)	.661(**)	.651(**)	.496(**)	.538(**)	.776(**)	.242(**)	.115	.600(**)	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.360	.000	
	N	117	196	235	252	232	235	238	65	238	255

** Correlation is significant at the 0.01 level (2-tailed).

Table 6
Organic Carbon Correlations/All Valid Data 2003-2005

		Douglas Pass	Ripple Creek Pass	Canyonlands NP	Great Sand Dunes NP	Mesa Verde NP	Mt. Zirkel WA	Rocky Mountain NP	Shamrock Mines	Weminuche WA	White River
Douglas Pass	Pearson Correlation	1	.924(**)	.929(**)	.894(**)	.867(**)	.928(**)	.683(**)	.198	.913(**)	.964(**)
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.197	.000	.000
	N	225	202	122	120	121	111	122	44	120	117
Ripple Creek Pass	Pearson Correlation	.924(**)	1	.888(**)	.836(**)	.643(**)	.961(**)	.762(**)	.593(**)	.850(**)	.903(**)
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	202	299	181	201	178	182	184	52	185	196
Canyonlands NP	Pearson Correlation	.929(**)	.888(**)	1	.794(**)	.712(**)	.874(**)	.639(**)	.441(**)	.847(**)	.895(**)
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.001	.000	.000
	N	122	181	241	238	233	221	236	55	236	235
Great Sand Dunes NP	Pearson Correlation	.894(**)	.836(**)	.794(**)	1	.666(**)	.826(**)	.581(**)	.740(**)	.894(**)	.850(**)
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000
	N	120	201	238	258	235	238	241	64	241	252
Mesa Verde NP	Pearson Correlation	.867(**)	.643(**)	.712(**)	.666(**)	1	.652(**)	.553(**)	.672(**)	.757(**)	.690(**)
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000
	N	121	178	233	235	238	222	233	53	233	232
Mt. Zirkel WA	Pearson Correlation	.928(**)	.961(**)	.874(**)	.826(**)	.652(**)	1	.788(**)	.689(**)	.834(**)	.900(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000
	N	111	182	221	238	222	241	224	57	225	235
Rocky Mountain NP	Pearson Correlation	.683(**)	.762(**)	.639(**)	.581(**)	.553(**)	.788(**)	1	.694(**)	.581(**)	.625(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.000
	N	122	184	236	241	233	224	243	55	239	238
Shamrock Mines	Pearson Correlation	.198	.593(**)	.441(**)	.740(**)	.672(**)	.689(**)	.694(**)	1	.362(**)	.706(**)
	Sig. (2-tailed)	.197	.000	.001	.000	.000	.000	.000		.008	.000
	N	44	52	55	64	53	57	55	67	52	65
Weminuche WA	Pearson Correlation	.913(**)	.850(**)	.847(**)	.894(**)	.757(**)	.834(**)	.581(**)	.362(**)	1	.900(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.008		.000
	N	120	185	236	241	233	225	239	52	243	238
White River	Pearson Correlation	.964(**)	.903(**)	.895(**)	.850(**)	.690(**)	.900(**)	.625(**)	.706(**)	.900(**)	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	N	117	196	235	252	232	235	238	65	238	255

** Correlation is significant at the 0.01 level (2-tailed).

Table 7
Fine Mass Correlations/All Valid Data 2003-2005

		Douglas Pass	Ripple Creek Pass	Canyonlands NP	Great Sand Dunes NP	Mesa Verde NP	Mt. Zirkel WA	Rocky Mountain NP	Shamrock Mines	Weminuche WA	White River
Douglas Pass	Pearson Correlation	1	.918(**)	.691(**)	.107	.296(**)	.291(**)	.142(*)	.073	.683(**)	.373(**)
	Sig. (2-tailed)		.000	.000	.115	.000	.000	.035	.394	.000	.000
	N	234	208	224	216	217	209	220	140	223	220
Ripple Creek Pass	Pearson Correlation	.918(**)	1	.711(**)	.651(**)	.704(**)	.880(**)	.610(**)	.789(**)	.579(**)	.811(**)
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	208	300	272	286	260	269	272	143	275	288
Canyonlands NP	Pearson Correlation	.691(**)	.711(**)	1	.461(**)	.537(**)	.590(**)	.527(**)	.381(**)	.471(**)	.626(**)
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000	.000	.000
	N	224	272	339	325	317	310	326	149	329	331
Great Sand Dunes NP	Pearson Correlation	.107	.651(**)	.461(**)	1	.672(**)	.634(**)	.508(**)	.835(**)	.305(**)	.648(**)
	Sig. (2-tailed)	.115	.000	.000		.000	.000	.000	.000	.000	.000
	N	216	286	325	348	311	320	323	148	326	340
Mesa Verde NP	Pearson Correlation	.296(**)	.704(**)	.537(**)	.672(**)	1	.698(**)	.543(**)	.857(**)	.512(**)	.708(**)
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000
	N	217	260	317	311	326	301	313	140	317	317
Mt. Zirkel WA	Pearson Correlation	.291(**)	.880(**)	.590(**)	.634(**)	.698(**)	1	.699(**)	.866(**)	.517(**)	.822(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000
	N	209	269	310	320	301	334	309	144	314	325
Rocky Mountain NP	Pearson Correlation	.142(*)	.610(**)	.527(**)	.508(**)	.543(**)	.699(**)	1	.627(**)	.341(**)	.826(**)
	Sig. (2-tailed)	.035	.000	.000	.000	.000	.000		.000	.000	.000
	N	220	272	326	323	313	309	337	143	328	329
Shamrock Mines	Pearson Correlation	.073	.789(**)	.381(**)	.835(**)	.857(**)	.866(**)	.627(**)	1	.266(**)	.849(**)
	Sig. (2-tailed)	.394	.000	.000	.000	.000	.000	.000		.001	.000
	N	140	143	149	148	140	144	143	162	145	158
Weminuche WA	Pearson Correlation	.683(**)	.579(**)	.471(**)	.305(**)	.512(**)	.517(**)	.341(**)	.266(**)	1	.843(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.001		.000
	N	223	275	329	326	317	314	328	145	340	332
White River	Pearson Correlation	.373(**)	.811(**)	.626(**)	.648(**)	.708(**)	.822(**)	.826(**)	.849(**)	.843(**)	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	N	220	288	331	340	317	325	329	158	332	354

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Table 8
Coarse Mass Correlations/All Valid Data 2003-2005

		Douglas Pass	Ripple Creek Pass	Canyonlands NP	Great Sand Dunes NP	Mesa Verde NP	Mt. Zirkel WA	Rocky Mountain NP	Shamrock Mines	Weminuche WA	White River
Douglas Pass	Pearson Correlation	1	.881(**)	.358(**)	.234(**)	.053	.030	-.007	.335(**)	.049	.461(**)
	Sig. (2-tailed)		.000	.000	.000	.416	.642	.912	.000	.453	.000
	N	241	210	240	239	239	239	239	159	239	238
Ripple Creek Pass	Pearson Correlation	.881(**)	1	.371(**)	.263(**)	.037	-.054	-.015	.449(**)	.042	.556(**)
	Sig. (2-tailed)	.000		.000	.000	.523	.351	.794	.000	.474	.000
	N	210	306	293	306	293	304	293	161	293	305
Canyonlands NP	Pearson Correlation	.358(**)	.371(**)	1	.094	.016	-.011	.007	.038	.095	.213(**)
	Sig. (2-tailed)	.000	.000		.073	.754	.832	.893	.617	.069	.000
	N	240	293	367	366	366	366	366	174	366	365
Great Sand Dunes NP	Pearson Correlation	.234(**)	.263(**)	.094	1	-.082	.003	-.003	-.063	.004	.159(**)
	Sig. (2-tailed)	.000	.000	.073		.117	.959	.951	.392	.933	.002
	N	239	306	366	380	366	378	366	187	366	378
Mesa Verde NP	Pearson Correlation	.053	.037	.016	-.082	1	-.155(**)	-.004	.121	-.096	.012
	Sig. (2-tailed)	.416	.523	.754	.117		.003	.933	.111	.066	.822
	N	239	293	366	366	366	366	366	174	366	365
Mt. Zirkel WA	Pearson Correlation	.030	-.054	-.011	.003	-.155(**)	1	-.001	-.081	-.001	.019
	Sig. (2-tailed)	.642	.351	.832	.959	.003		.980	.269	.979	.712
	N	239	304	366	378	366	378	366	186	366	377
Rocky Mountain NP	Pearson Correlation	-.007	-.015	.007	-.003	-.004	-.001	1	.582(**)	.007	-.005
	Sig. (2-tailed)	.912	.794	.893	.951	.933	.980		.000	.894	.922
	N	239	293	366	366	366	366	366	174	366	365
Shamrock Mines	Pearson Correlation	.335(**)	.449(**)	.038	-.063	.121	-.081	.582(**)	1	.071	.051
	Sig. (2-tailed)	.000	.000	.617	.392	.111	.269	.000		.355	.485
	N	159	161	174	187	174	186	174	187	174	186
Weminuche WA	Pearson Correlation	.049	.042	.095	.004	-.096	-.001	.007	.071	1	.011
	Sig. (2-tailed)	.453	.474	.069	.933	.066	.979	.894	.355		.838
	N	239	293	366	366	366	366	366	174	366	365
White River	Pearson Correlation	.461(**)	.556(**)	.213(**)	.159(**)	.012	.019	-.005	.051	.011	1
	Sig. (2-tailed)	.000	.000	.000	.002	.822	.712	.922	.485	.838	
	N	238	305	365	378	365	377	365	186	365	378

** Correlation is significant at the 0.01 level (2-tailed).

Back Trajectories. The National Oceanic and Atmospheric Administration (NOAA) Air Resources Laboratory (ARL) HYbrid Single-Particle Lagrangian Integrated Trajectory model or HYSPLIT has been used to calculate back trajectories for IMROVE sample days in 2005 at the Douglas Pass and Ripple Creek monitoring sites. A total of 576 back trajectory points have been calculated for each sample day at each site. These points represent 48 hours of back trajectory points for every two-hour period during a sample day. This spread of points provides a comprehensive picture of the possible air mass source areas for the sample during the previous two days. With about 120 sample days in 2005, this exercise yields about 70,000 back trajectory points for each monitoring site.

The 2005 back trajectory points for Ripple Creek are shown in Figure 1. The distribution of points shows that the air arriving at Ripple Creek typically passes through western and northwestern Colorado. Utah, Wyoming, northwestern New Mexico and Arizona are key source regions for this site. There was very little transport from eastern Colorado during this period.

The back trajectory points for those samples with calculated extinctions within the worst 20 percentile for 2005 are shown in Figure 2. Since all paths to Ripple Creek must travel through the counties in the vicinity of the site, the highest density of points is near Ripple Creek. The trajectory points also suggest that the Salt Lake City and Four Corners areas may be source regions for the samples with the worst calculated extinction levels.

It is possible to estimate the average air quality contributions of any area covered by back trajectory points using a moving spatial average method. This approach eliminates the central tendency bias near the monitor caused by the convergence of all trajectories at this point. The use of the HYSPLIT model and the EDAS40 input meteorological data set are key steps in the process.

The meteorological data that drives the HYSPLIT model has a grid resolution of 40 kilometers. While this grid resolution can lead to trajectory location errors, particularly in areas of complex terrain, the input weather data and HYSPLIT have incorporated adjustments that improve the mean accuracy of calculated back trajectories. The input meteorological data is a hybrid of modeled and measured data, and it has been heavily weighted by actual observations from a very large number of surface stations, aircraft sensors, and other sources of real time weather information. HYSPLIT also employs routines for interpolating the data to a resolution much finer than 40 kilometers. While individual HYSPLIT back trajectories can suffer from errors associated with poor matches between the HYSPLIT wind fields and actual wind fields, it seems that on average the back trajectories provide a good representation of air mass source regions. When applied to ozone source regions along Colorado's Front Range, HYSPLIT back trajectories and the moving spatial average accurately located source regions within urban area valleys known to have persistent cold pooling and emissions trapping (see reference 1). There appeared to be a high degree of correspondence between local terrain features that affect emissions and the analyzed contributions of these areas.

For this analysis, the extinction or concentration values associated with a given sample were assigned to each of the calculated back trajectory points for that sample. A moving spatial average was applied to all of the point concentrations or extinction values for the

entire year. A computer routine placed an oval window with a height of 0.4 degrees latitude and width of 0.4 degrees longitude over each grid node (the grid node spacing was 0.2 degrees latitude in the Y direction and 0.2 degrees longitude in the X direction). If the number of points in the oval was 130 or higher, the routine calculated the mean extinction or concentration for that node. The distribution of mean values on the map was then contoured using the Kriging method. Limiting the averaging to nodes with a high number of points reduces the chances that a few points associated with a high or low sample concentration or extinction bias the estimates of source contributions from that locale. The resulting map provides an estimate of the mean contribution for any area covered by the analysis.

The results for Ripple Creek calculated extinction (Bext) are presented in Figure 3. Contours were not generated for nodes or averaging windows with too few points, and this causes the irregular distribution of contoured areas. This map suggests that the regions that contributed to the highest average extinction values were generally some distance from the site. Both the Four Corners region and an area that may include sources in southern Idaho, southwestern Wyoming, and the Salt Lake City area seem to be significant source areas for higher extinction.

The moving spatial average method was also applied to the trajectory points for days that account for the worst 20-percentile calculated extinction. The spatial threshold for averaging was lowered to 30 points in this case. The results shown in Figure 4 highlight a corridor from northwest Colorado to southwestern Wyoming, the Salt Lake City area, and southern Idaho. While the results in Figures 3 and 4 could benefit from the inclusion of data for other years, they suggest that distant source regions were significant for high extinction days.

The 2005 back trajectory points for Douglas Pass are shown in Figure 5. The distribution of points shows that the air arriving at Douglas Pass typically passes through the western one-fifth of Colorado. Utah, Wyoming, northwestern New Mexico and Arizona are key source regions for this site. There was very little transport from eastern Colorado during this period.

The back trajectory points for those Douglas Pass samples with calculated extinctions within the worst 20 percentile for 2005 are shown in Figure 6. Since all paths to Douglas Pass must travel through the counties in the vicinity of the site, the highest density of points is near Douglas Pass. The trajectory points also suggest that central and eastern Utah, Arizona, and the Four Corners areas may be source regions for the samples with the worst calculated extinction levels.

The results for the spatial moving average source-region analysis for calculated extinction (Bext) at Douglas Pass are presented in Figure 7. Contours were not generated for nodes or averaging windows with too few points, and this explains the irregular distribution of the contoured areas. This map suggests that the regions that contributed to the highest average extinction values included the Four Corners region, southwestern Colorado in general, eastern Utah, and Mesa County in Colorado. A similar spatial analysis of positive matrix factorization results described later in this section sheds some light on emissions sources that may be responsible for these “hotspots”.

The moving spatial average method was also applied to the trajectory points for days that account for the worst 20-percentile calculated extinction. The spatial threshold for averaging was lowered to 30 points in this case. The results shown in Figure 8 suggest that the highest source regions are in southwestern Colorado, northwestern New Mexico, and eastern Utah. While the results in Figures 3 and 4 are subject to further review and could benefit from the inclusion of data for other years, they suggest that these source regions were significant for high extinction days.

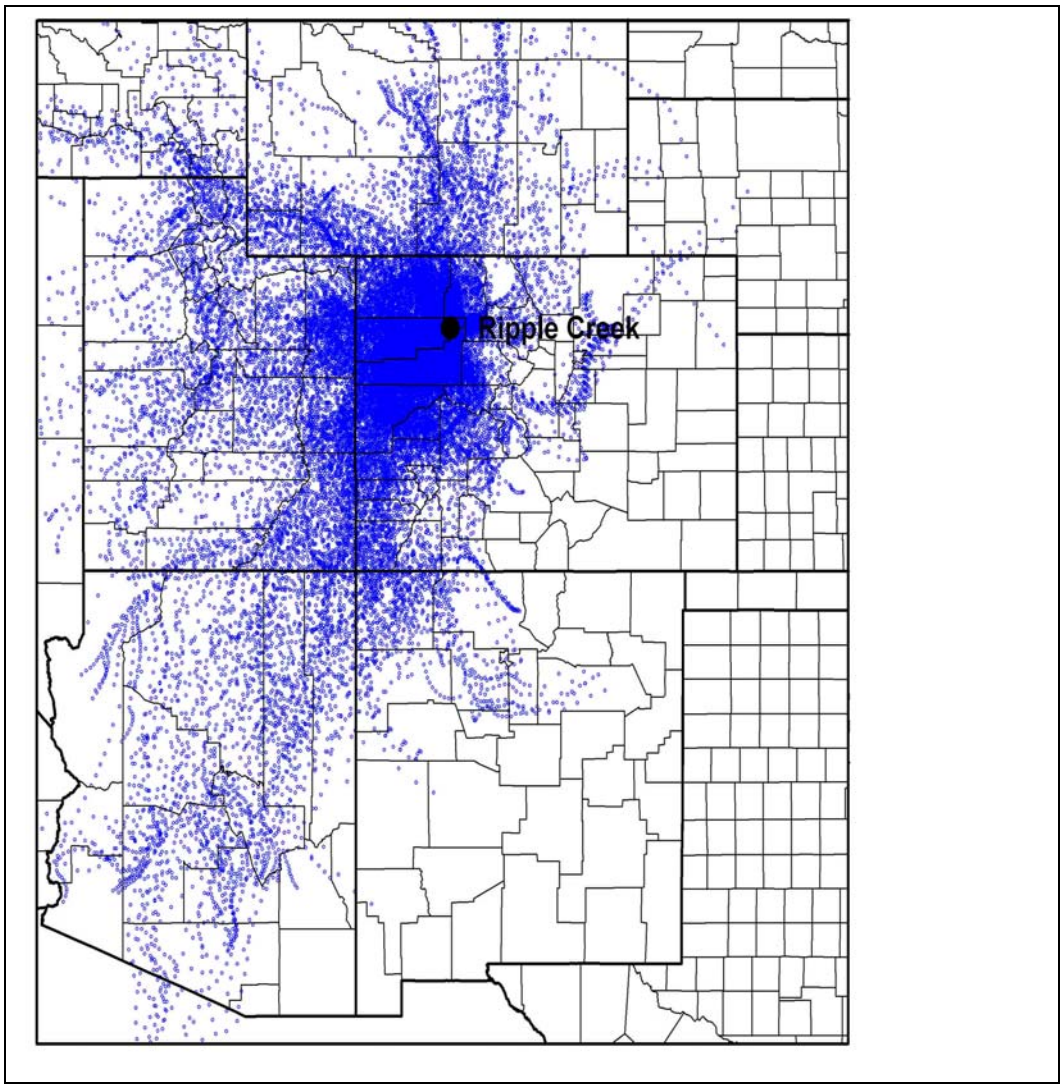

Figure 1. NOAA ARL HYSPLIT back trajectory points for Ripple Creek for roughly 120 IMPROVE sample days in 2005. A total of 576 back trajectory points have been calculated for each sample day. These 576 points represent 48 hours of back trajectory points for every two-hour period during a sample day.

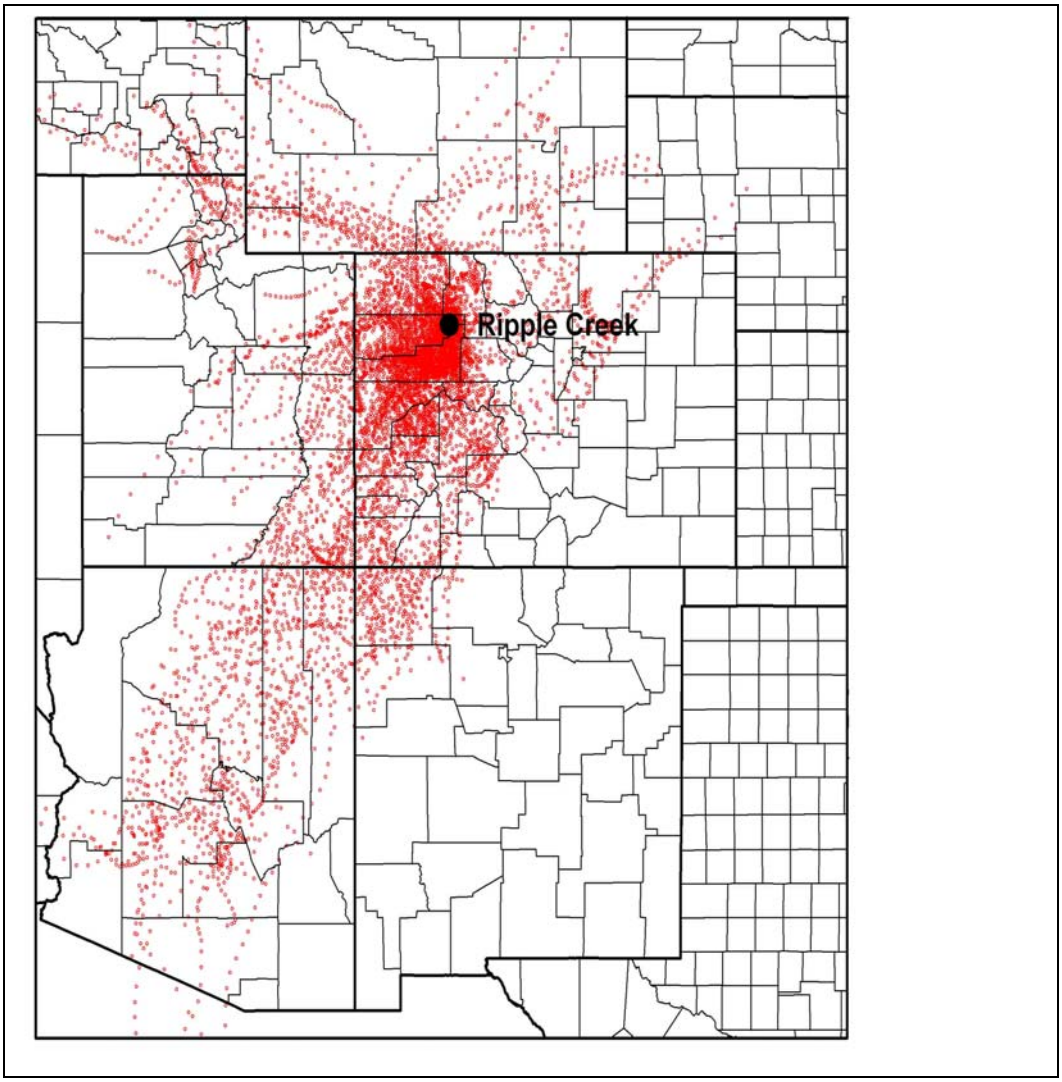

Figure 2. NOAA ARL HYSPLIT back trajectory points for Ripple Creek for the IMPROVE sample days with the worst 20 percentile calculated extinction in 2005. A total of 576 back trajectory points have been calculated for each worst-case day. These points represent 48 hours of back trajectory points for every two-hour period during a sample day.

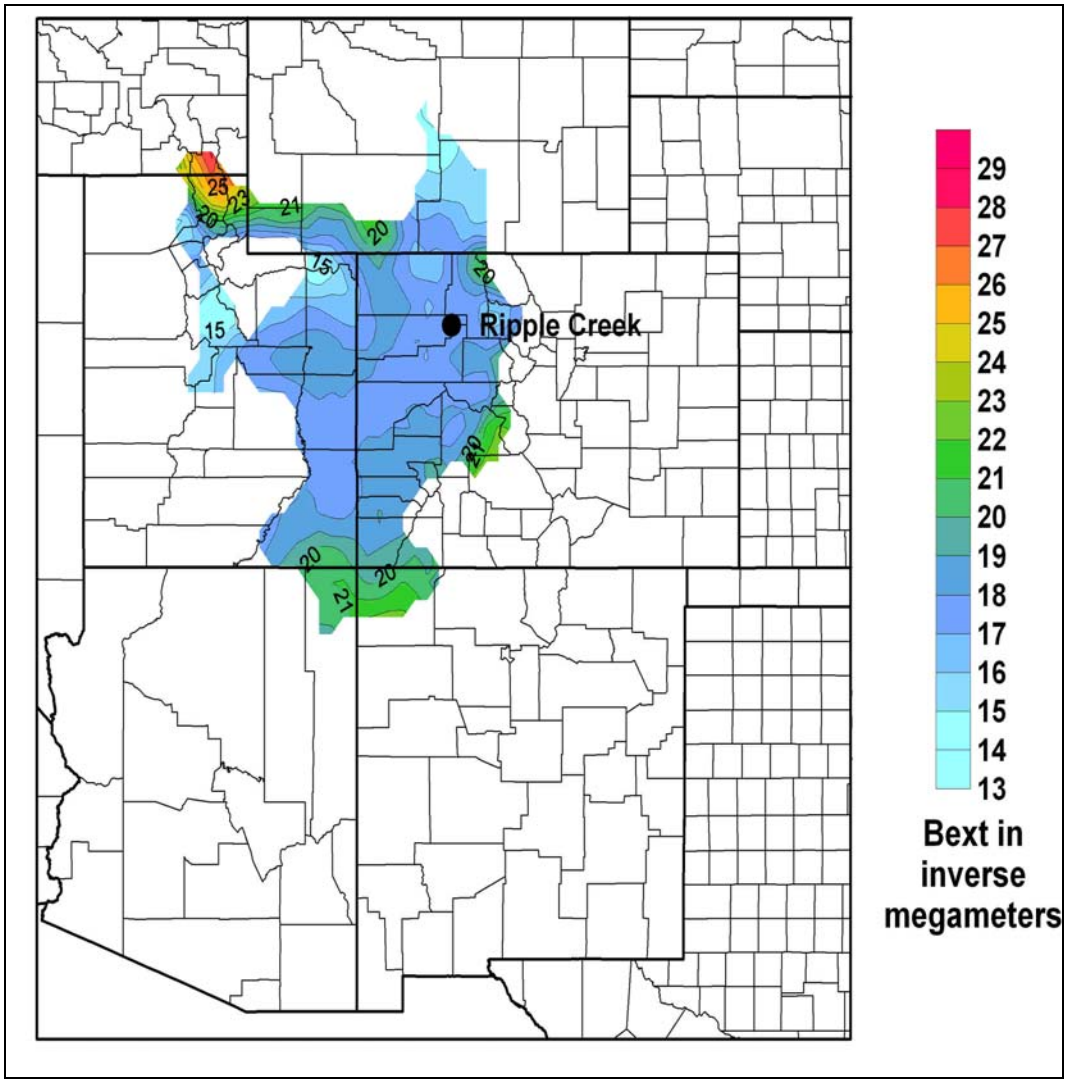

Figure 3. A map of the mean calculated extinction at Ripple Creek in 2005 when the air mass is from a specific region (based on a spatial moving average analysis of extinction values associated with HYSPLIT back trajectory points). The highest source region is in northern Utah and southern Idaho. When an air mass is from this region, for example, the mean IMPROVE calculated extinction is above 20 inverse megameters.

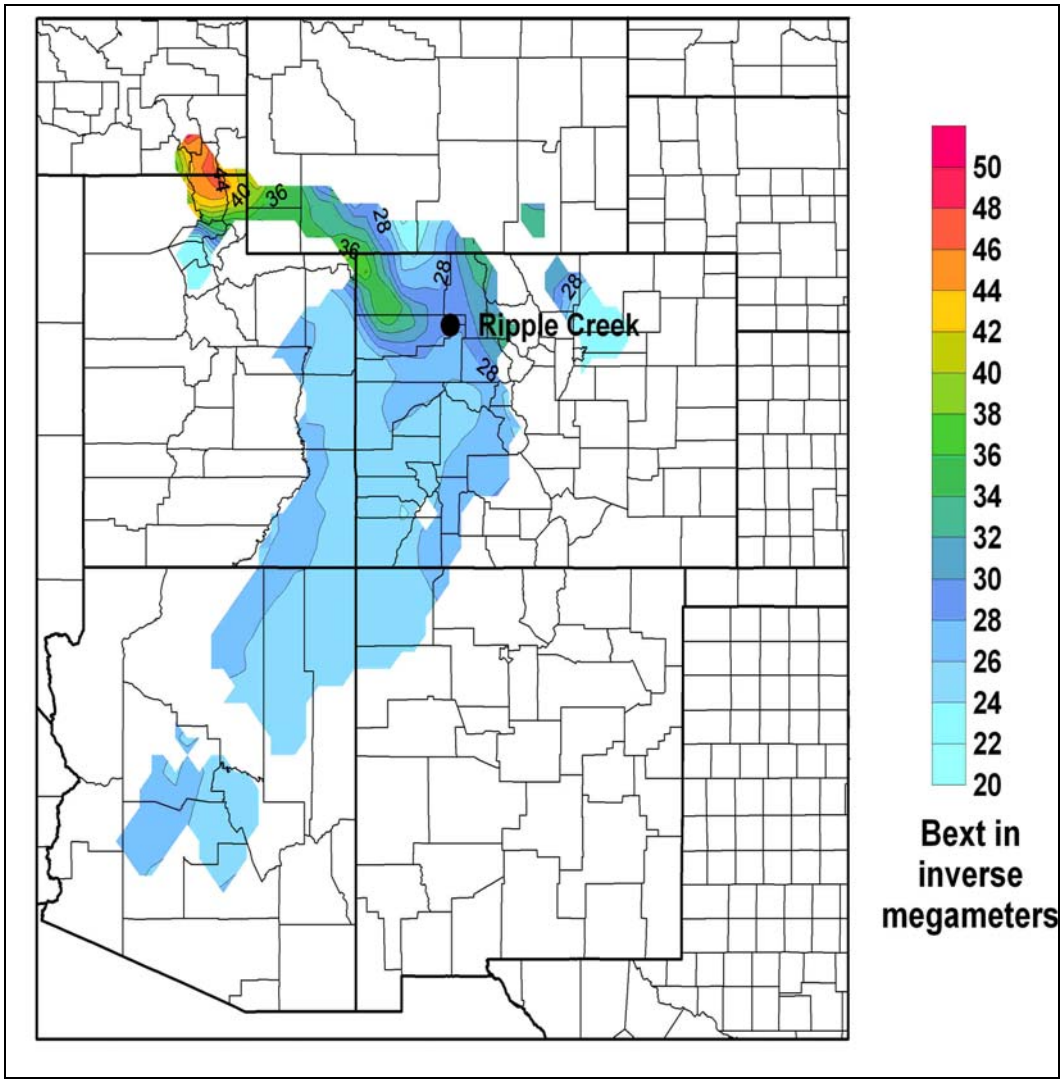

Figure 4. A map of the mean calculated extinction at Ripple Creek in 2005 for the worst 20 percentile sample days when the air mass is from a specific region (based on a spatial moving average analysis of extinction values associated with HYSPLIT back trajectory points). The highest source region is in northern Utah and southern Idaho. When an air mass is from this region, for example, the mean IMPROVE calculated extinction during the worst 20 percentile events is above 40 inverse megameters.

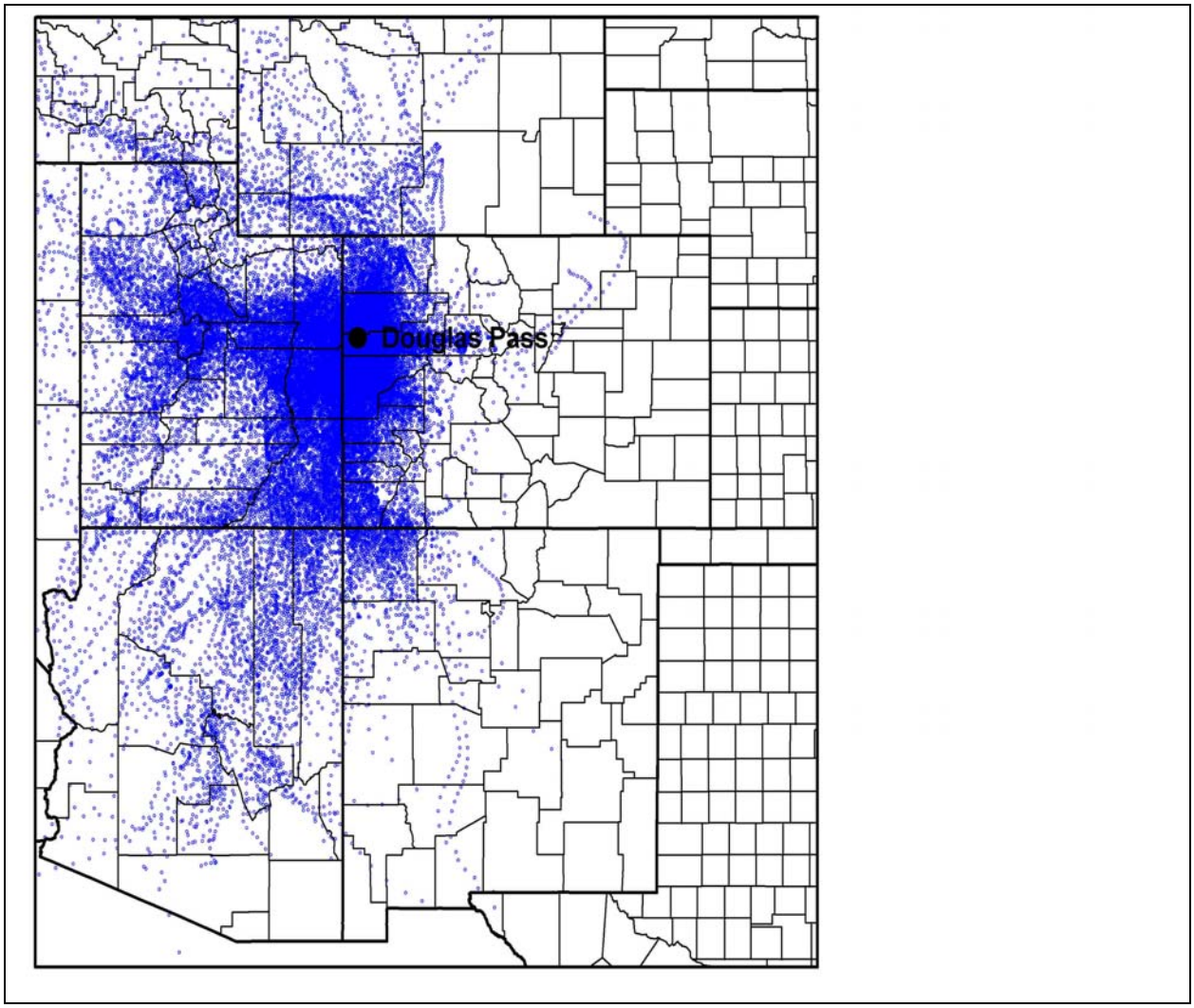

Figure 5. NOAA ARL HYSPLIT back trajectory points for Douglas Pass for roughly 120 IMPROVE sample days in 2005. A total of 576 back trajectory points have been calculated for each sample day. These 576 points represent 48 hours of back trajectory points for every two-hour period during a sample day.

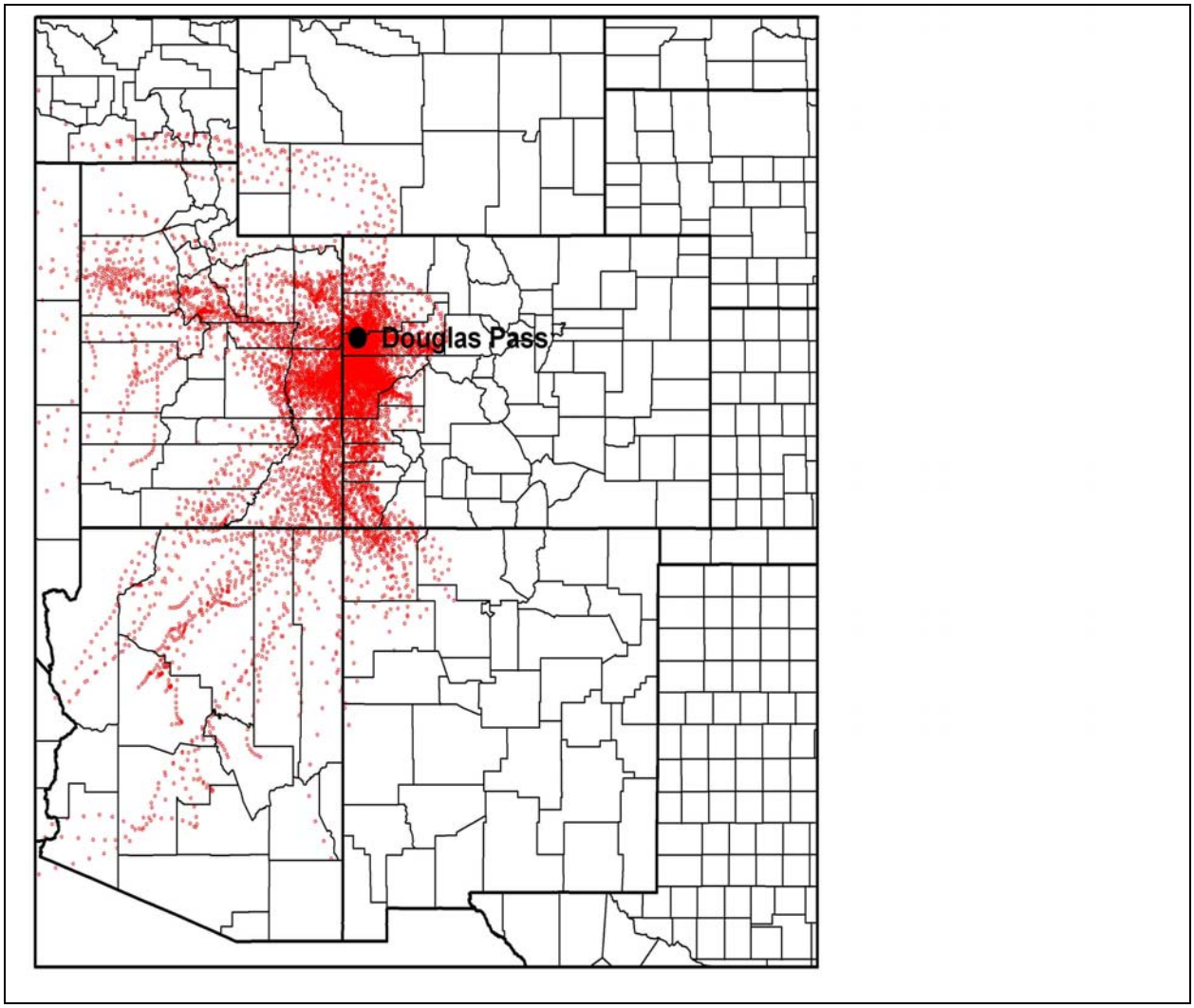

Figure 6. NOAA ARL HYSPLIT back trajectory points for Douglas Pass for the IMPROVE sample days with the worst 20 percentile calculated extinction in 2005. A total of 576 back trajectory points have been calculated for each worst-case day. These points represent 48 hours of back trajectory points for every two-hour period during a sample day.

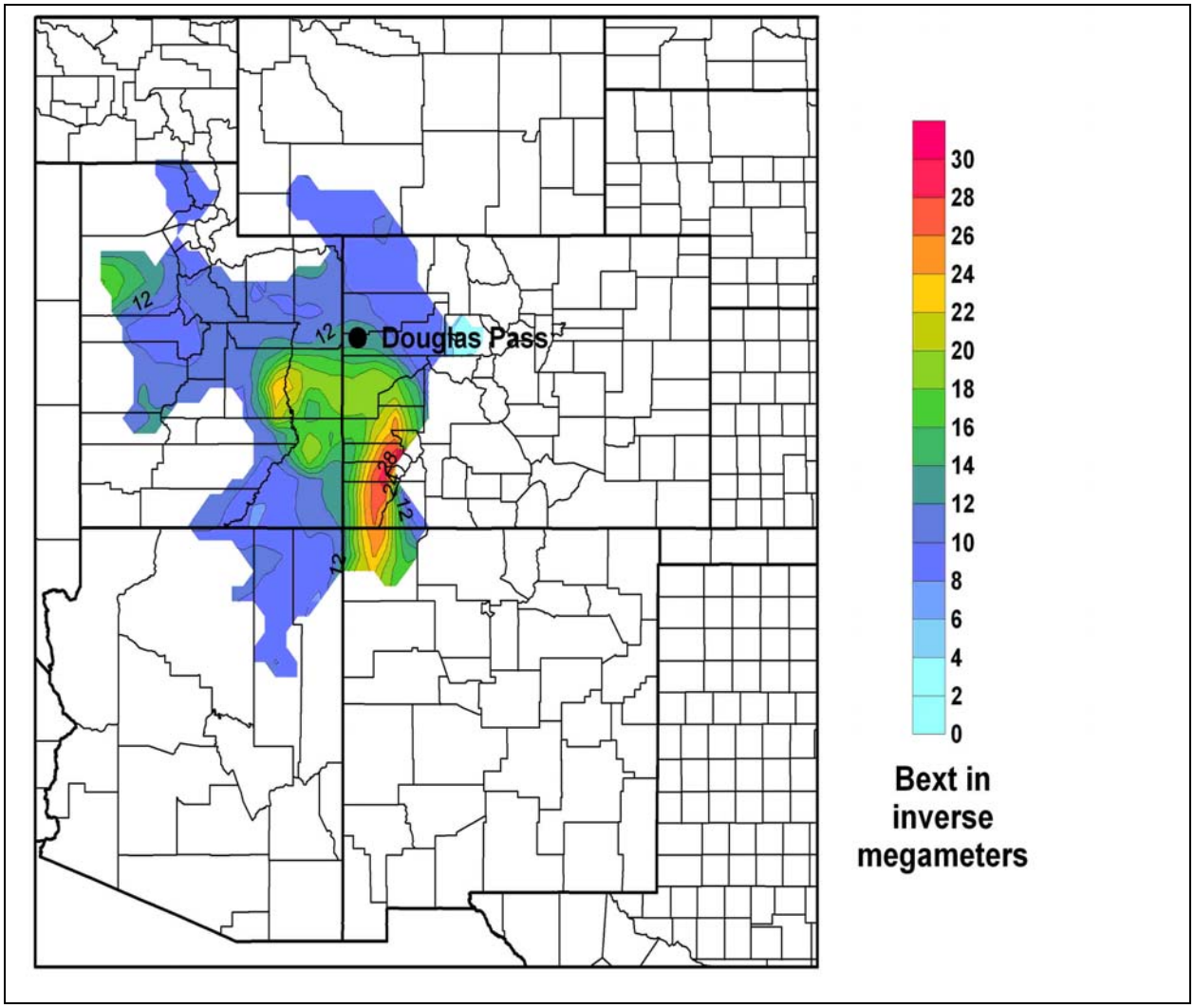

Figure 7. A map of the mean calculated extinction at Douglas Pass in 2005 when the air mass is from a specific region (based on a spatial moving average analysis of extinction values associated with HYSPLIT back trajectory points). The highest source areas are in eastern Utah, southwestern Colorado, and northwestern New Mexico. When an air mass is from these regions, for example, the mean IMPROVE calculated extinction is above 22 inverse megameters.

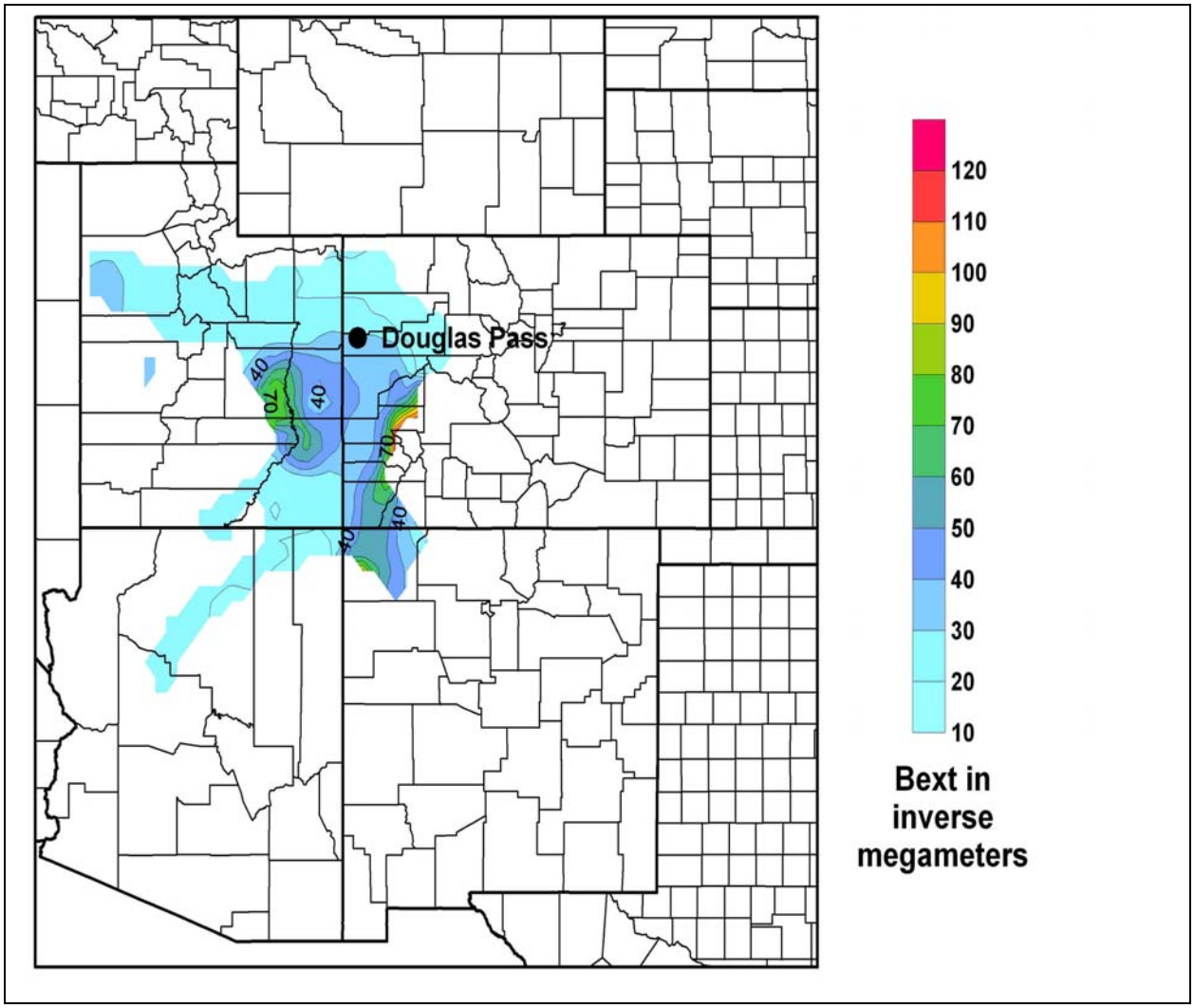

Figure 8. A map of the mean calculated extinction at Douglas Pass in 2005 for the worst 20 percentile sample days when the air mass is from a specific region (based on a spatial moving average analysis of extinction values associated with HYSPLIT back trajectory points). The highest source region is in southwestern Colorado, northwestern New Mexico, and eastern Utah. When an air mass is from these regions, for example, the mean IMPROVE calculated extinction during the worst 20 percentile events is above 60 inverse megameters.

Positive Matrix Factorization. Positive Matrix Factorization (PMF) has been applied to IMPROVE sample data for Douglas Pass and Ripple Creek for 2004 and 2005. PMF is an analytical procedure that can be used to identify air pollution source categories and to estimate the contributions of these source categories for individual air quality samples in a time series. In this sense it is similar to Chemical Mass Balance analysis or CMB, but emissions or source profiles are not needed in PMF. PMF, which is similar to factor analysis, identifies source profiles through statistical analyses. Unlike factor analysis or principal component analysis, PMF yields non-negative estimates of the contributions of each source category to the total measurement for each sample. In the analysis reported here, PMF has been applied to IMPROVE aerosol fine mass concentrations. A description of the method and discussion of the output diagnostics attached to the end of

this TSD.

Four source categories were identified for Douglas Pass. These categories are a sulfate group or source factor, a geologic material source factor, a wood combustion source factor, and a nitrate group or source factor. The final diagnostic for the analysis is shown in Figure 9. This is a linear regression of the sum of the concentrations from all four source categories against the measured or gravimetric fine-mass concentrations from the IMPROVE sample. The slope is near one at 0.97. This shows that on average the PMF results reproduce the total measured concentrations from the samples. The intercept is near zero, and the R-squared is high at 0.92.

The PMF results for Ripple Creek were not used. Although most of the diagnostics were good, the regression of total PMF concentrations against measured concentrations was weak. The slope was nearly one at 0.996, but the intercept was relatively high at 1.2 ug/m3, and the R-squared value was quite low at 0.34. There was too much scatter in the plot, suggesting that the modeled contributions would often provide a poor representation of reality.

Figure 10 presents the mean 2004-2005 concentrations (source profiles) for each category in ug/m3. Time series of source factor contributions for each sample day are presented in Figures 11 and 12 for 2004 and 2005, respectively.

The sulfate group probably represents power generation sources and probably also includes mobile sources and oil and gas combustion. The nitrate group probably represents power generation sources, mobile sources, and fuel combustion in general. There does not seem to be enough of a signal in the Douglas Pass data to separate these two groups into smaller categories. Model performance declined significantly when more than four categories were included. Both groups represent secondary aerosols.

The sulfate group shows weak seasonality with lower values in the winter and a few major episodes. The geologic group is lowest in the winter and highest in the spring, and this is consistent with the annual patterns for blowing dust. The wood combustion group is lower in the winter and higher in summer and late fall. This is consistent with the seasonal patterns for wildfire and controlled burning smoke. The nitrate group is highest in the late winter and spring, and there are a few major episodes in late 2005.

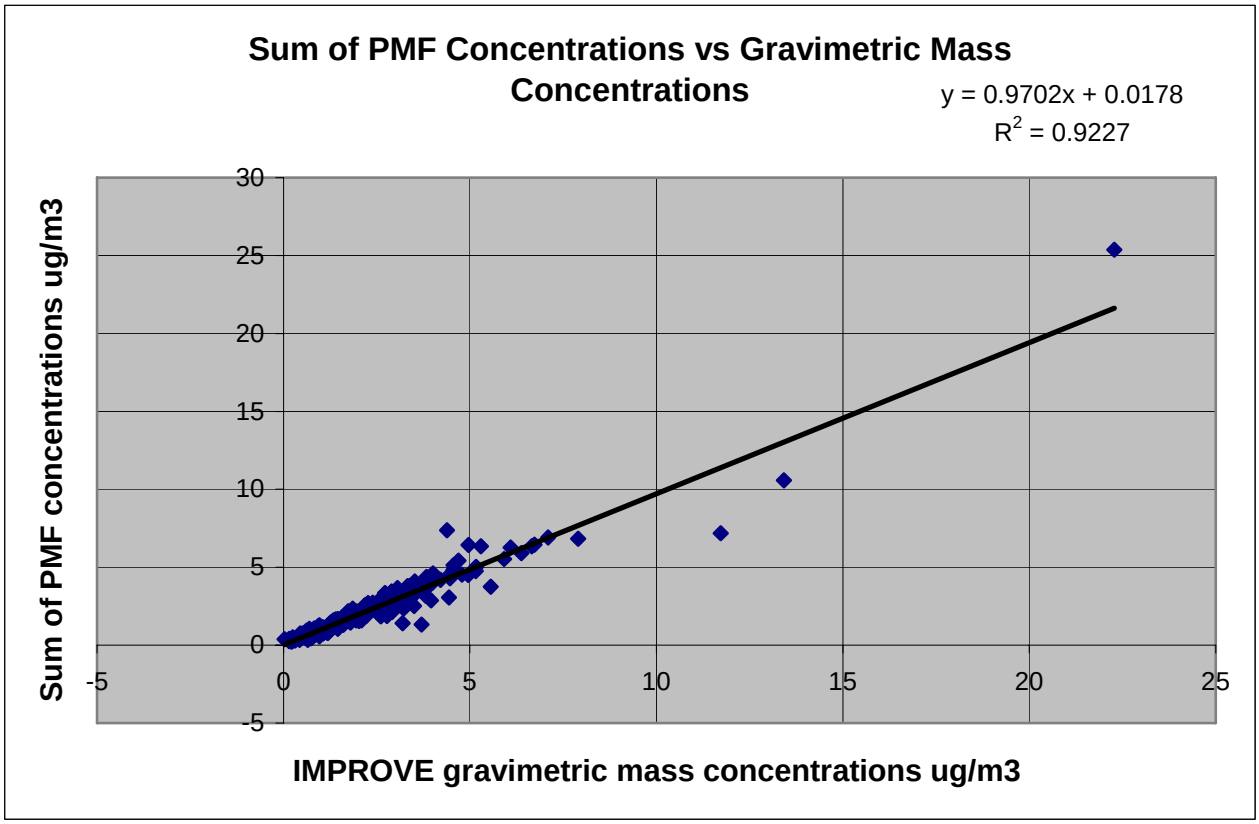

Figure 9. The sum of PMF concentrations versus gravimetric mass concentrations for Douglas Pass (2004-2005). The linear regression shows a nearly one-to-one correspondence and an R-squared of 0.92.

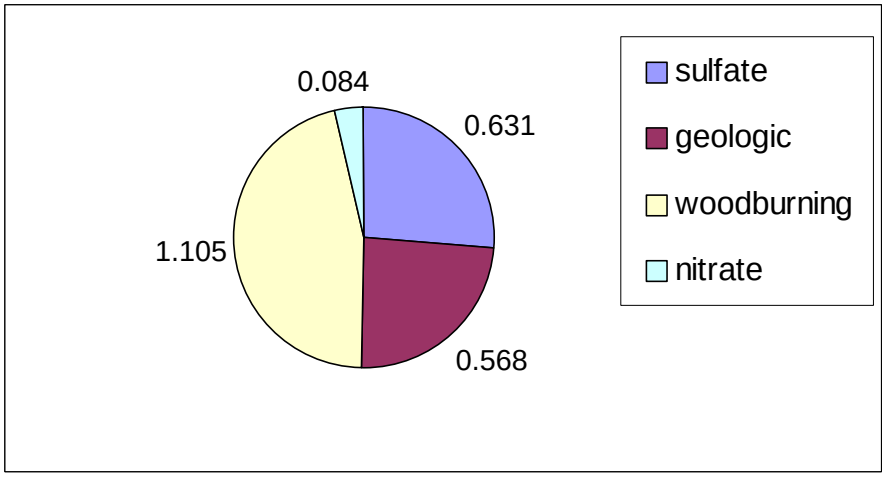

Figure 10. Mean 2004-2005 Douglas Pass concentrations (source profiles) for each category in ug/m3.

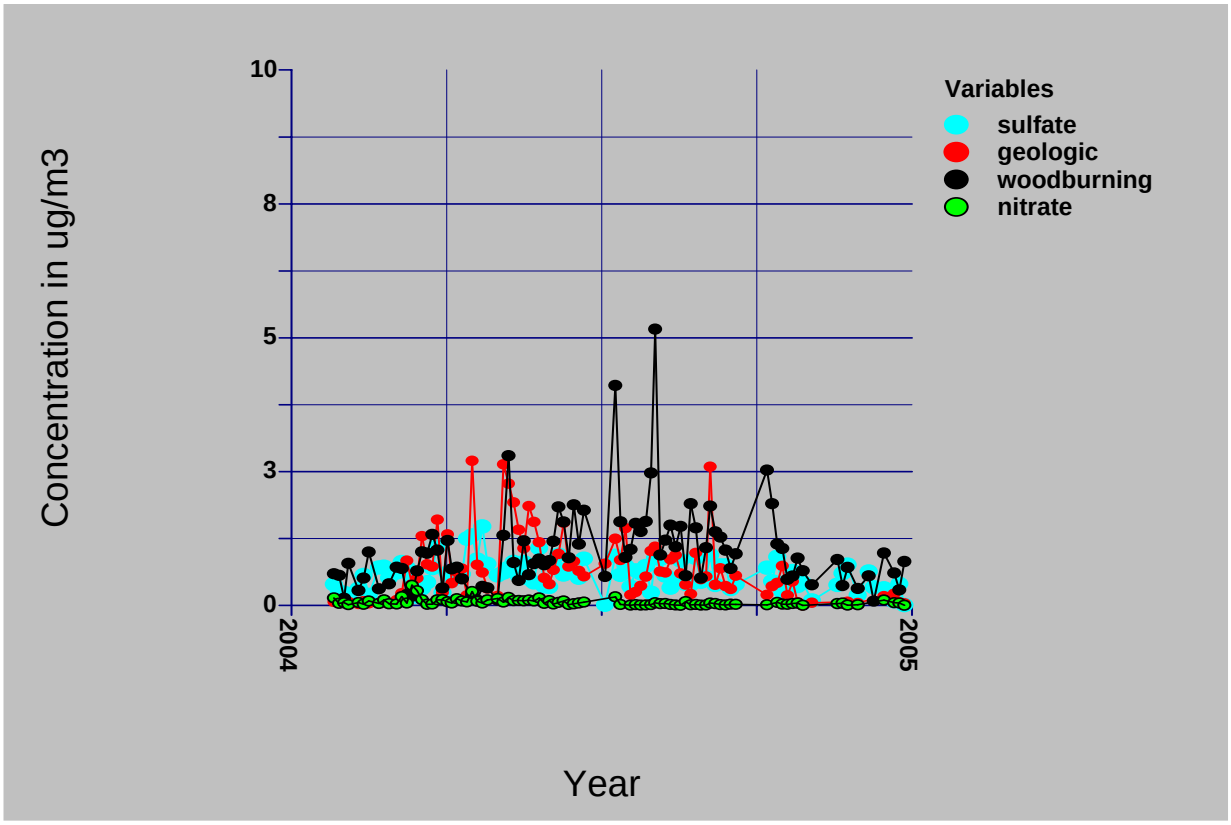

Figure 11. Time series of PMF source factor contributions at Douglas Pass in ug/m3 for each sample day for 2004.

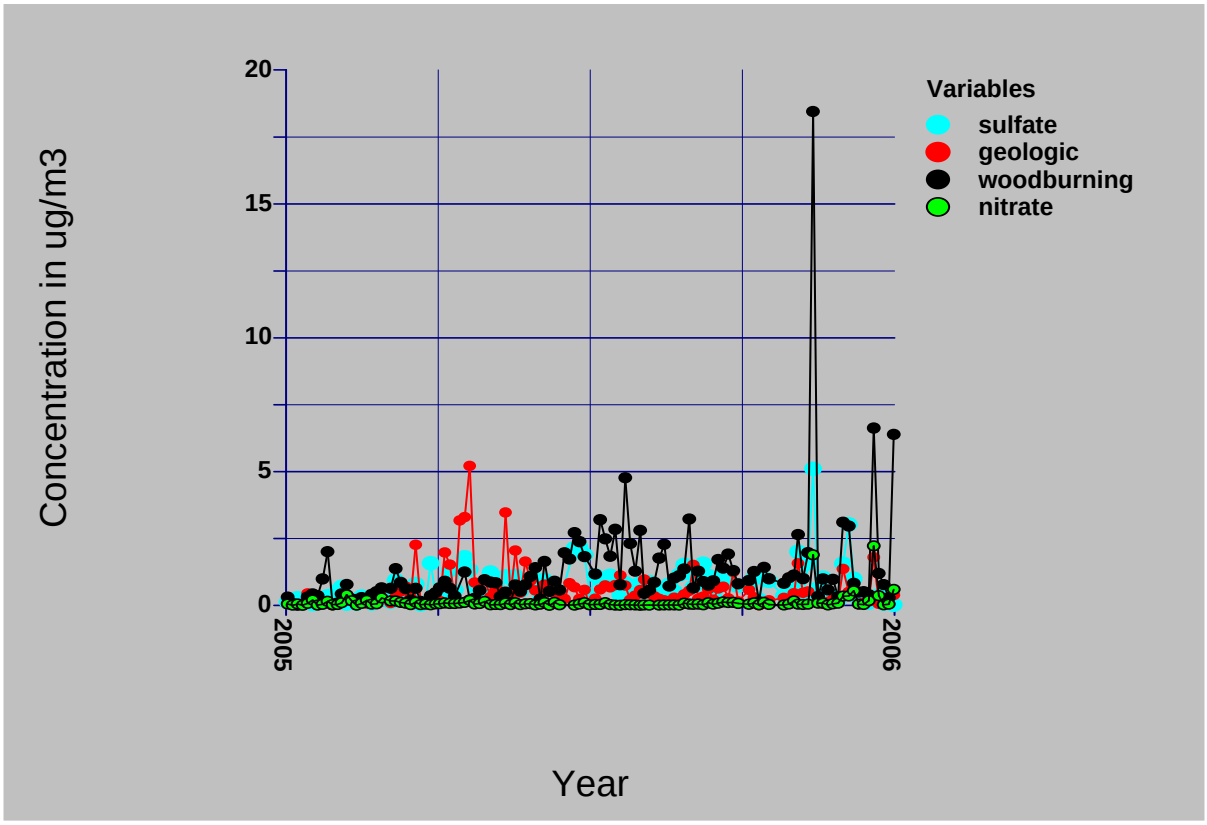

Figure 12. Time series of PMF source factor contributions at Douglas Pass in ug/m3 for each sample day for 2005.

The moving spatial average method was also applied to the Douglas Pass trajectory points associated with the sample-day contributions from each of the four PMF source groups for Douglas Pass for the year 2005. The results shown in Figure 13 highlight a sulfate group source corridor from northwest New Mexico northward through southwestern Colorado and another source area in eastern Utah. Contours for the geologic source factor are presented in Figure 14. These show significant source areas in the deserts of southern Utah and northern Arizona. These are known source areas for blowing dust NAAQS particulate matter exceedance events for fine particulates at Western Colorado monitoring sites. On average, transport from these areas results in a source group concentration of 1.2 ug/m3 or higher at Douglas Pass. Results for the wood combustion source factor are shown in Figure 15. The pattern for wood combustion is similar to the patterns for both the sulfate and nitrate source groups. As discussed in the Methods and Model Diagnostics section at the end of this TSD, this may in part be due to inadequate separation between the sulfate and wood combustion categories in the PMF analysis. The wood combustion category does show a small hotspot in Mesa County (the large corbel-shaped county south of Douglas Pass), which would be consistent with residential wood combustion and agricultural burning. Wildfires and controlled burns in eastern Utah and western Colorado may have influenced the overall pattern. An investigation of specific wildfire events and their possible effects on these results has not been completed. Finally, the results for the nitrate group are shown in Figure 16. Hotspot source areas are found in eastern Utah, southwestern Colorado, and over Mesa County.

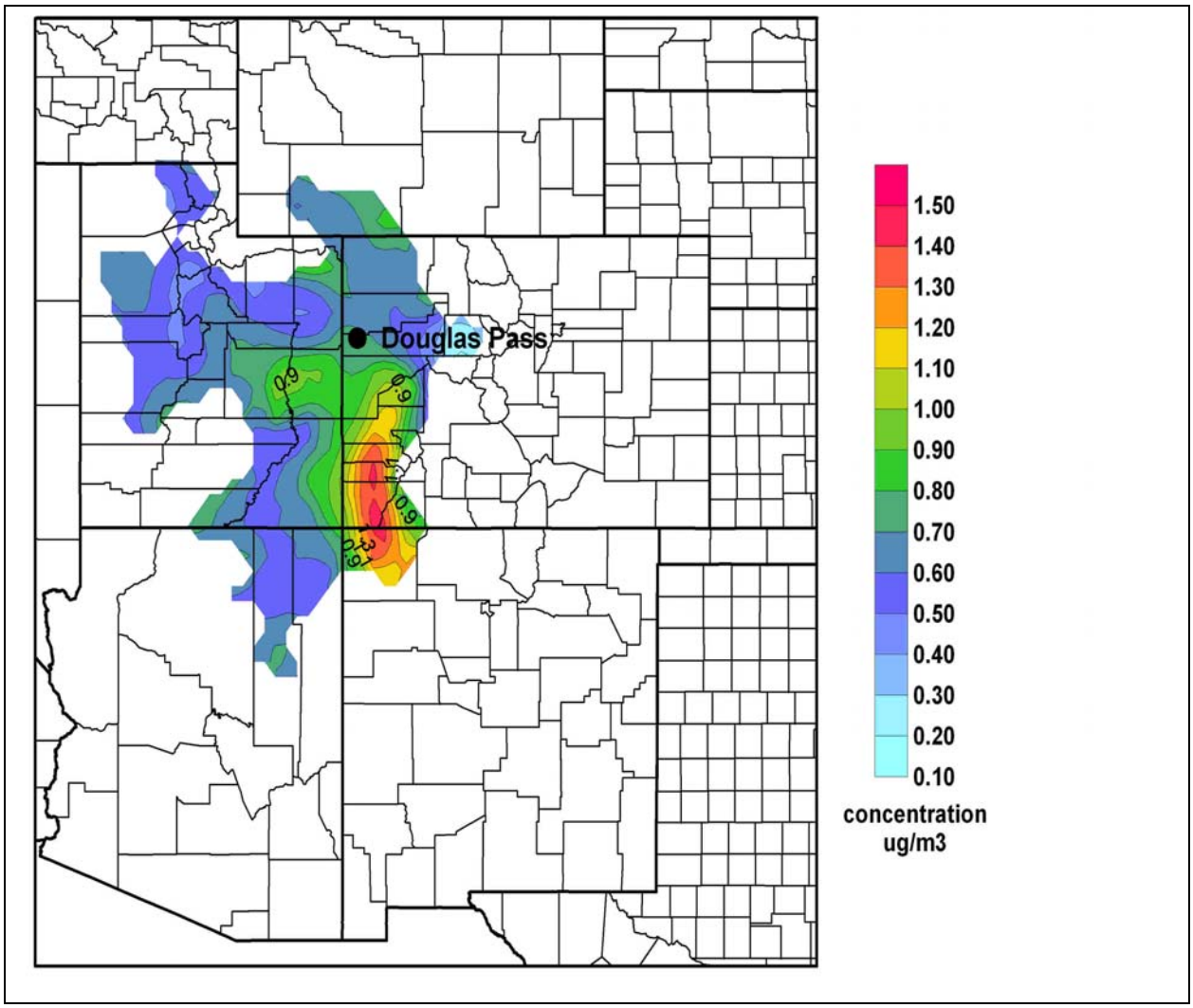

Figure 13. A map of the mean PMF sulfate group concentrations at Douglas Pass in 2005 when the air mass is from a specific region (based on a spatial moving average analysis of sulfate factor concentrations associated with HYSPLIT back trajectory points). The highest source region is in southwestern Colorado, northwestern New Mexico, and eastern Utah. When an air mass is from northwestern New Mexico or southwestern Colorado, the average sulfate group concentration at Douglas Pass is above 1 ug/m3.

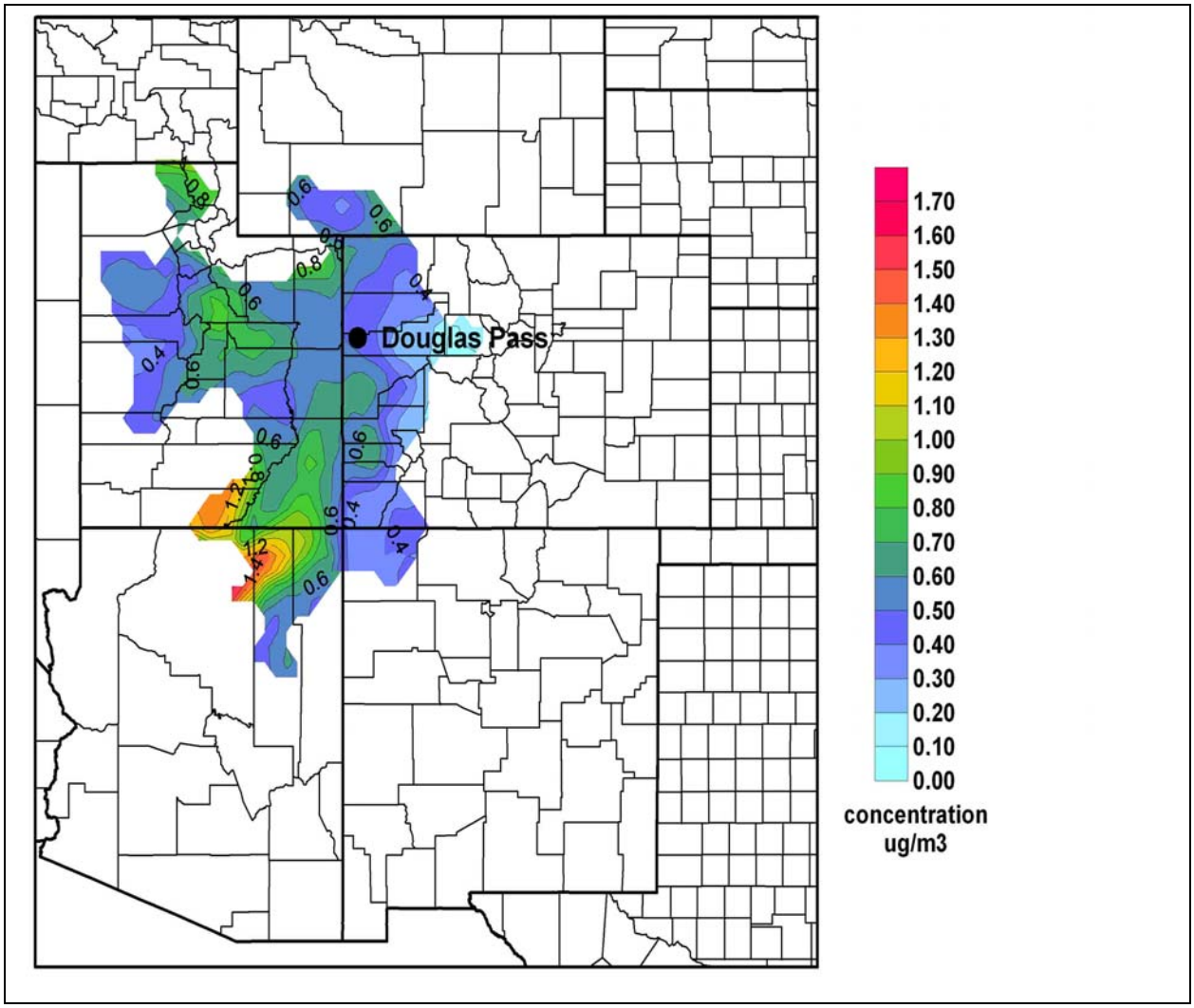

Figure 14. A map of the mean PMF geologic source category concentrations at Douglas Pass in 2005 when the air mass is from a specific region (based on a spatial moving average analysis of geologic factor concentrations associated with HYSPLIT back trajectory points). The highest source region is in southern Utah and northern Arizona. When an air mass is from these areas, the average geologic source category concentration at Douglas Pass is above 1 ug/m3.

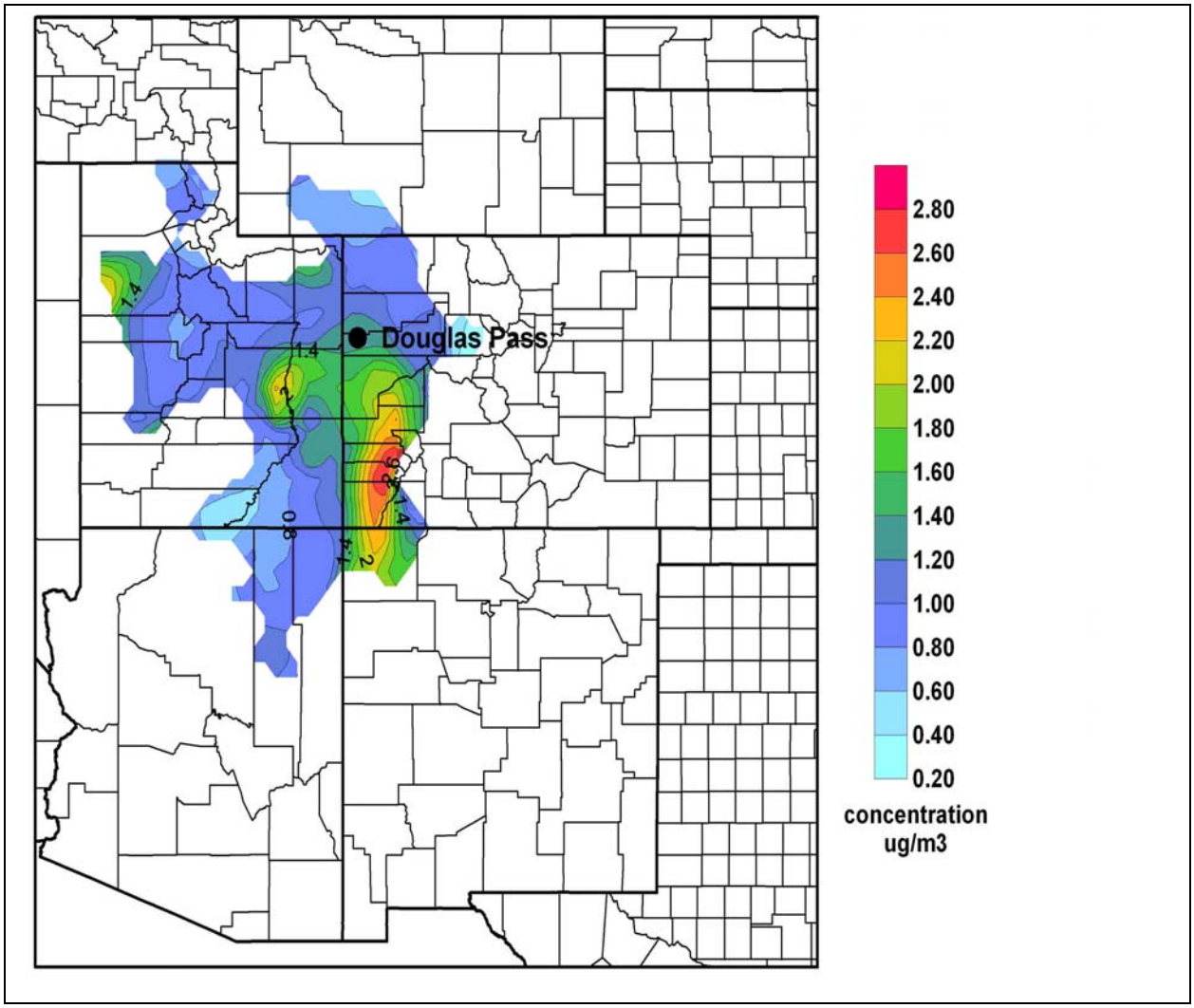

Figure 15. A map of the mean PMF wood combustion category concentrations at Douglas Pass in 2005 when the air mass is from a specific region (based on a spatial moving average analysis of wood combustion factor concentrations associated with HYSPLIT back trajectory points). The highest source regions are in northwestern New Mexico, southwestern Colorado, and eastern Utah. When an air mass is from these areas, the average wood combustion category concentration at Douglas Pass is above 2.0 ug/m3.

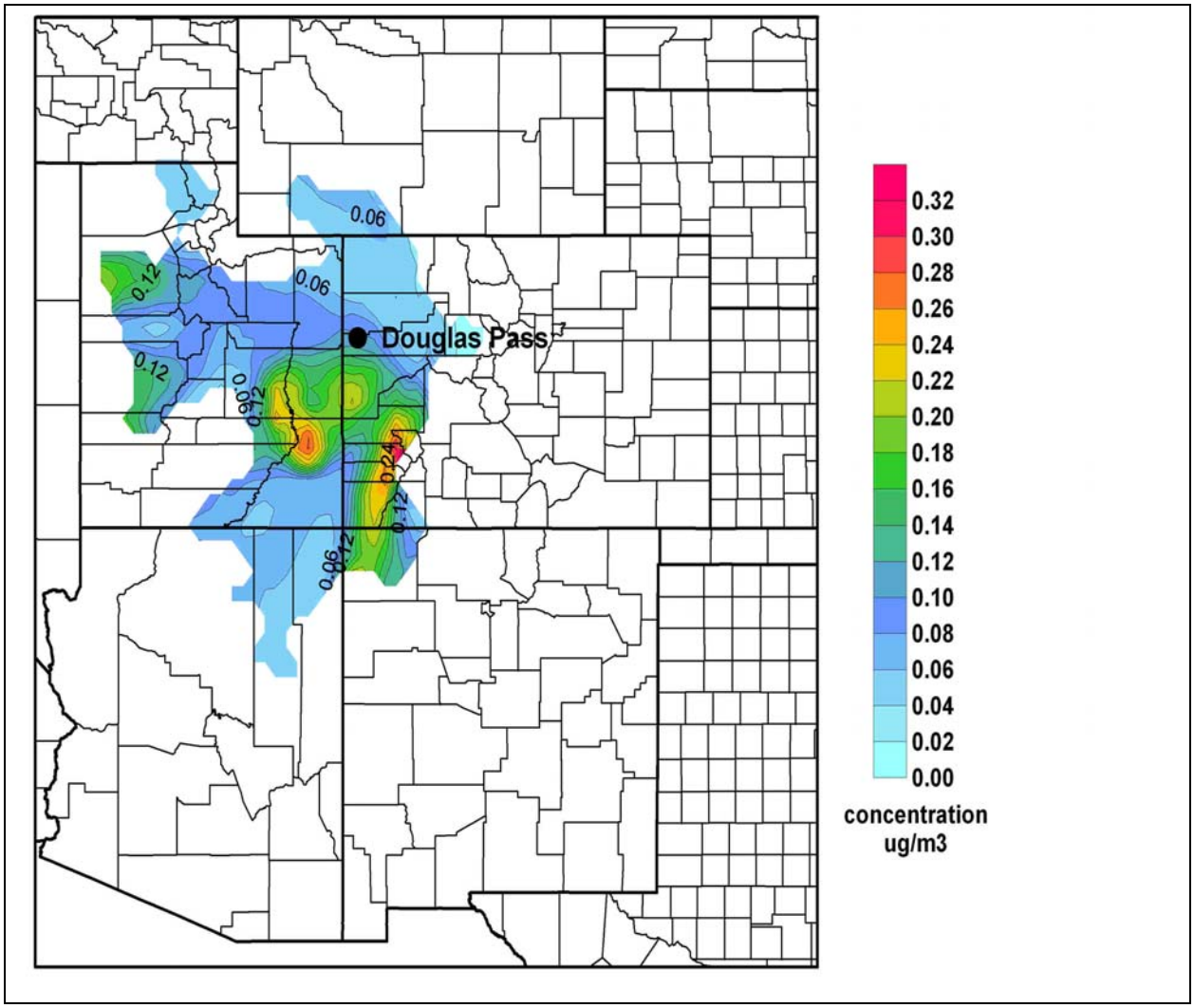

Figure 16. A map of the mean PMF nitrate group category concentrations at Douglas Pass in 2005 when the air mass is from a specific region (based on a spatial moving average analysis of nitrate factor concentrations associated with HYSPLIT back trajectory points). The highest source regions are in Mesa County (Colorado), southwestern Colorado, and eastern Utah. When an air mass is from these areas, the average wood combustion category concentration at Douglas Pass is above 0.2 ug/m3.

Point sources for SO2 and NOx are mapped in Figures 17 and 18, respectively. The diameters of the source locator symbols are roughly proportional to the sources' annual emissions.

Sulfur dioxide sources and source area contours for the Douglas Pass sulfate factor are shown in Figures 19 and 20. These clearly show that there is a correspondence between large SO2 sources and the analyzed source regions for the sulfate factor (based on back trajectories and monitoring at Douglas Pass). The source area for the highest concentrations is anchored to the Four Corners and San Juan power generation facilities in northwest New Mexico. It also lines up with oil and gas sources in the area and numerous smaller sources in southwestern Colorado. A source area in Utah is anchored to the Huntington power generation facility in east-central Utah. These three power

generation facilities appear to have had a major impact on worst-case concentrations in 2005. It is worth noting that sulfate is a secondary aerosol, and it often forms in areas downwind of sources. Source areas may sometimes be offset from point sources for SO2.

Nitrogen dioxide and nitrogen oxide sources and source area contours for the Douglas Pass nitrate factor are shown in Figures 21 and 22. One of the source areas for the highest concentrations is anchored to the Four Corners and San Juan power generation facilities in northwest New Mexico. There are also key source areas in Mesa County (Colorado) and east-central Utah. Nitrate is also a secondary aerosol, and it often forms in areas downwind of sources. Source areas may sometimes be offset from point sources for NOx and located in areas favorable for nitrate formation. Pooling of cold humid air in valleys can lead to accelerated nitrate formation as can emission of NOx into low clouds. There is not yet a clear explanation for the source areas showing up in east-Central Utah.

Yampa Valley sources do not seem to play a major role at either Ripple Creek or Douglas Pass. In some ways this is not surprising. Previous studies have demonstrated the role of Yampa Valley cold pooling and subsequent eastward transport of secondaries towards the Mt. Zirkel Wilderness Area. The major canyons and valleys in western Colorado can have a profound effect on local transport and can channel emissions away from some locations and towards others.

Reference:

1) Reddy, P. J., D. Wells, K. Briggs, C. Machovec, 2006. “Meteorological & climatological analysis of ozone back trajectories and their implications for the role of oil and gas field VOC emissions in Colorado and Front Range ozone formation”, Modeling, Meteorology, and Emissions Inventory Unit, Technical Services Program, Air Pollution Control Division, Colorado Department of Public Health and Environment.

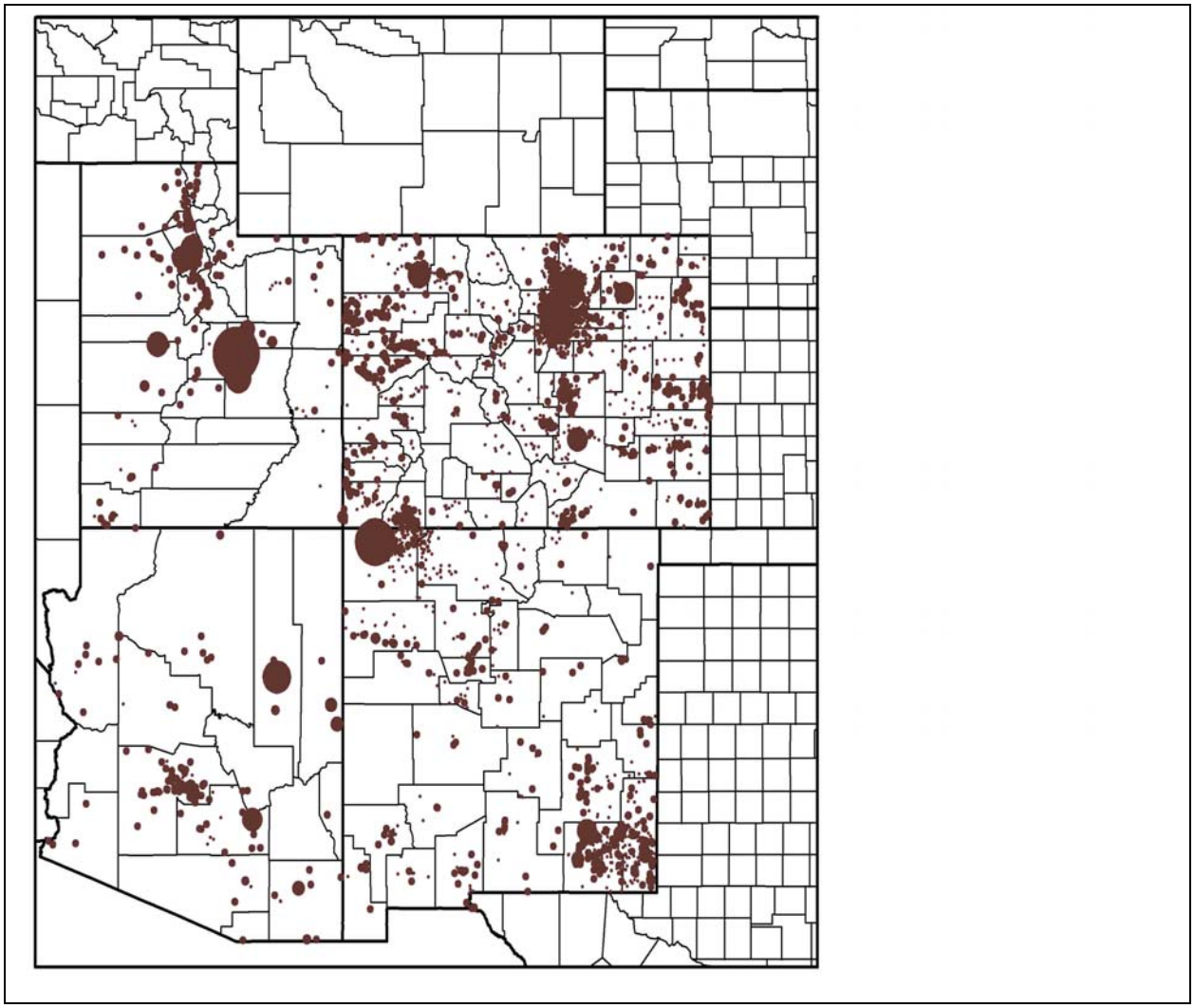

Figure 17. Sulfur dioxide point sources. The diameters of the source locator symbols are roughly proportional to the sources' annual emissions.

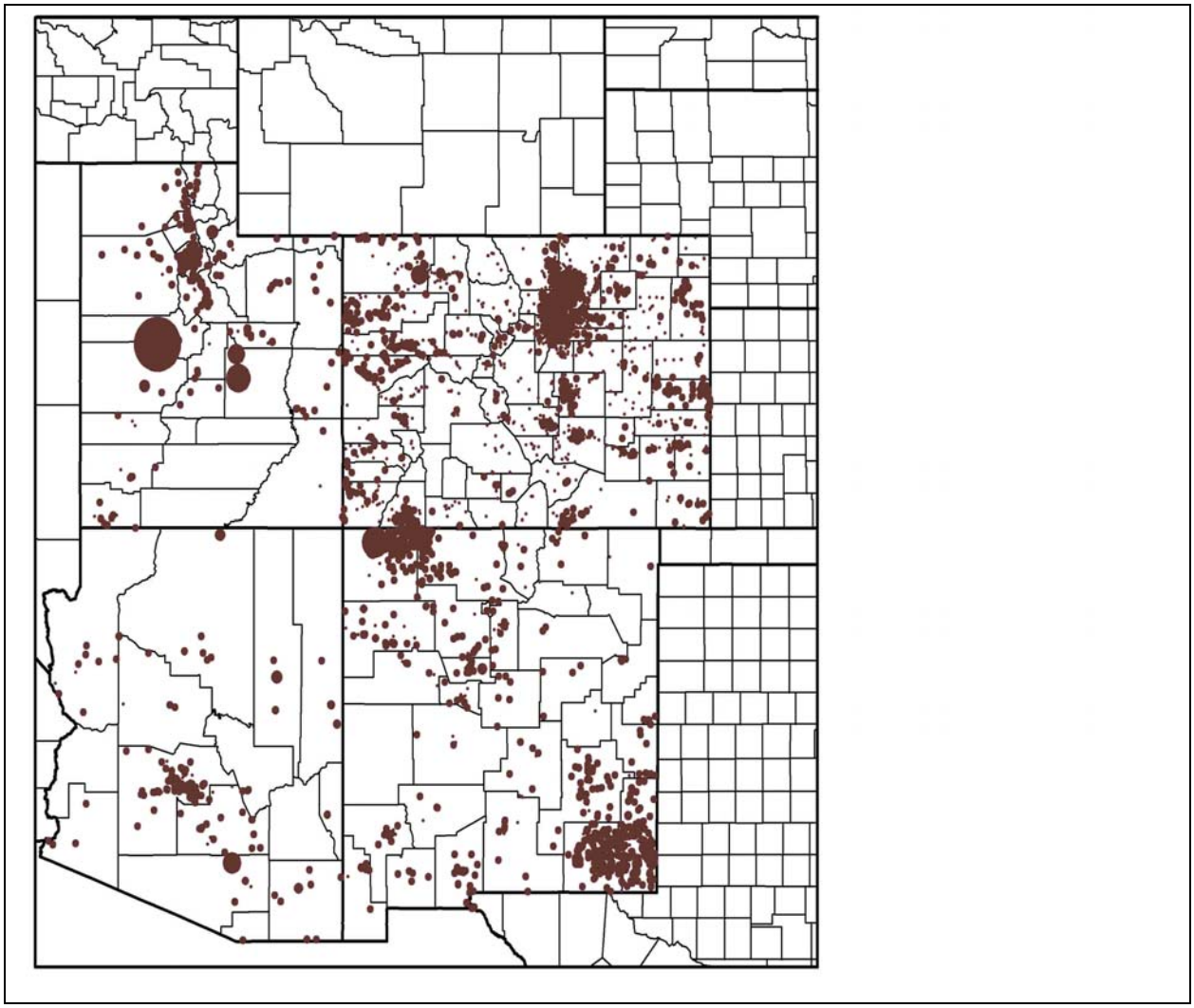

Figure 18. NOx point sources. The diameters of the source locator symbols are roughly proportional to the sources' annual emissions.

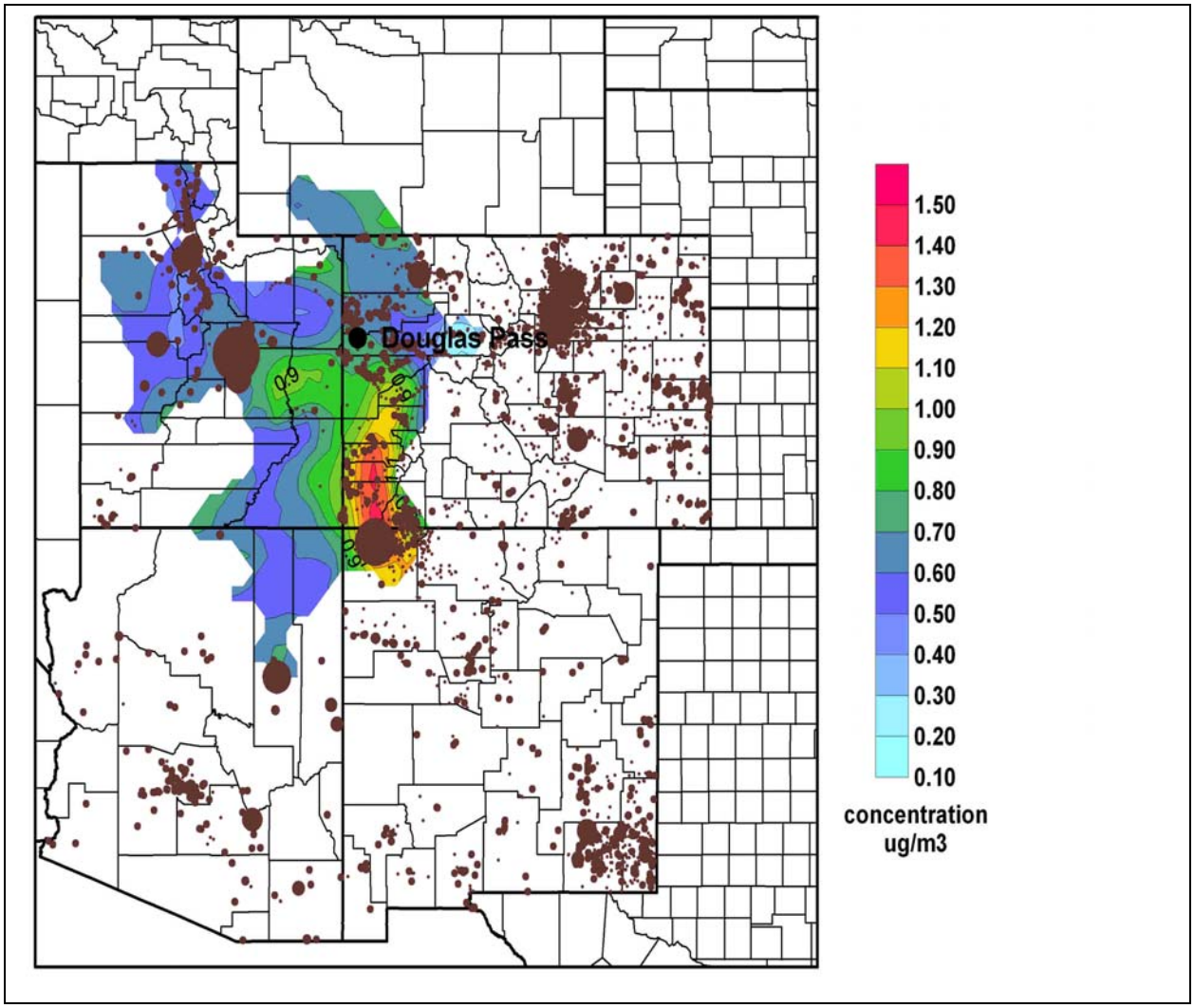

Figure 19. SO2 point sources, scaled by the magnitude of annual emissions, and contours of the Douglas Pass sulfate factor source areas from Figure 13 (based on a spatial moving average analysis of sulfate factor concentrations associated with HYSPLIT back trajectory points).

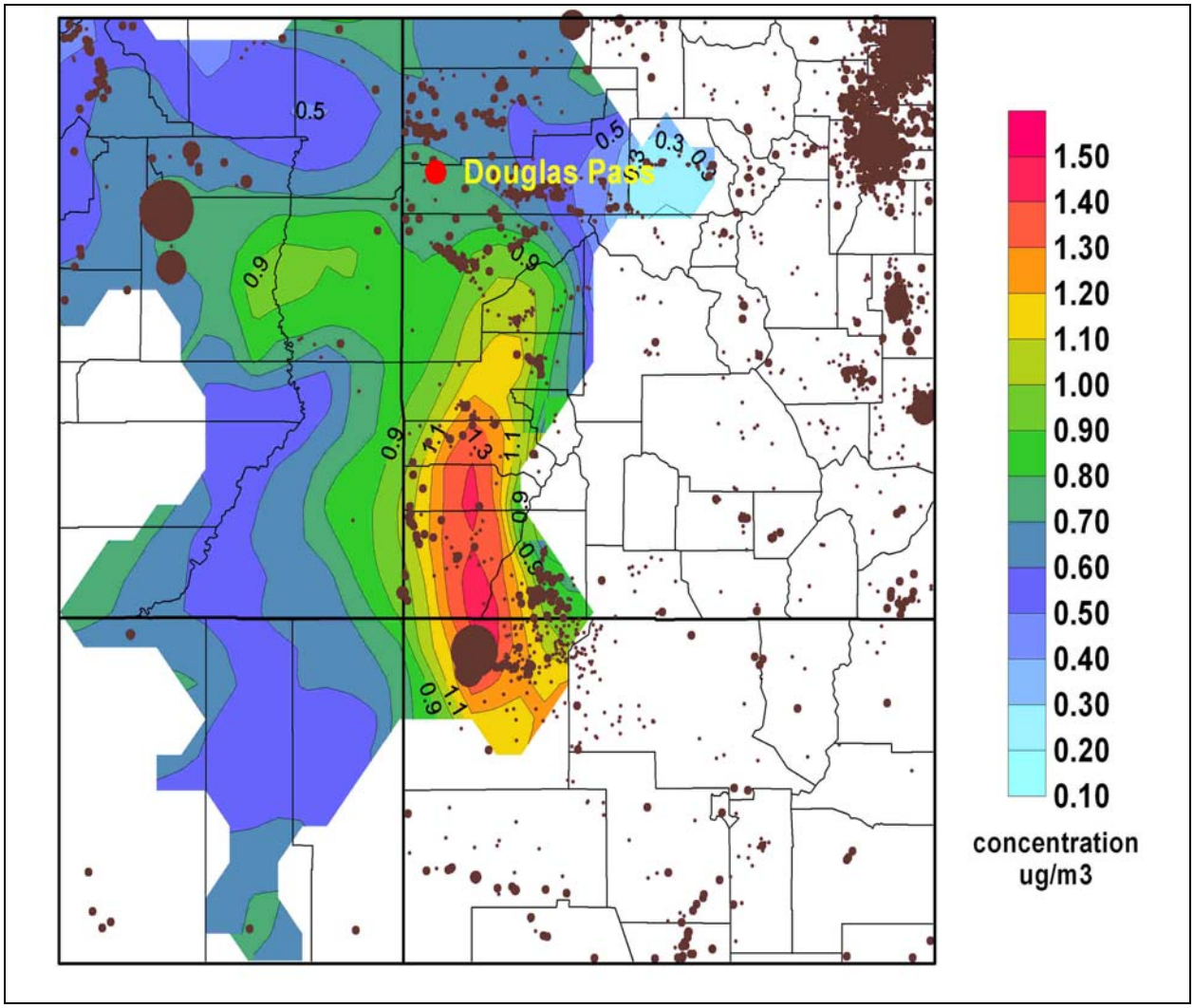

Figure 20. SO2 point sources, scaled by the magnitude of annual emissions, and contours of the Douglas Pass sulfate factor source areas from Figure 13 (based on a spatial moving average analysis of sulfate factor concentrations associated with HYSPLIT back trajectory points).

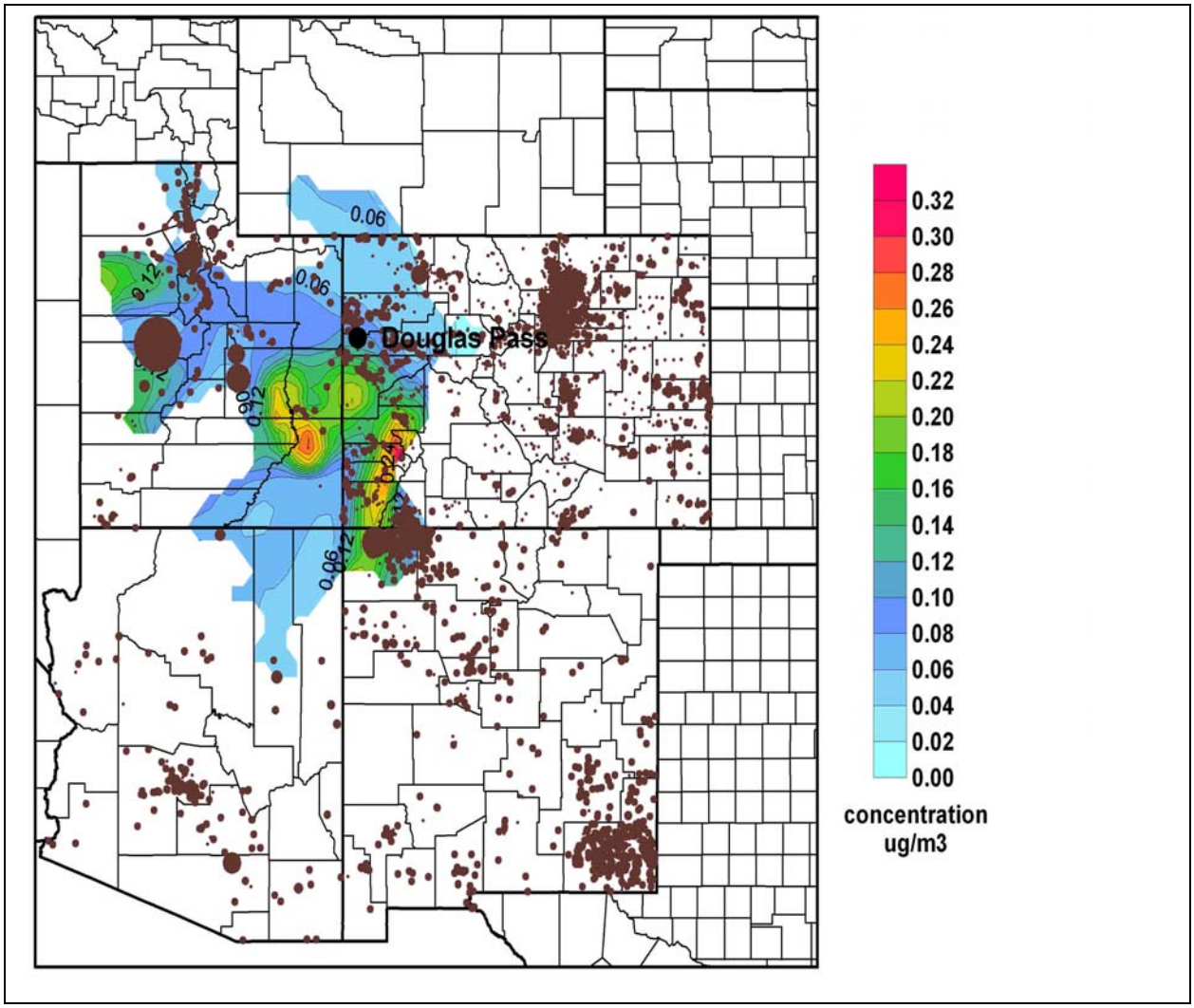

Figure 21. NOx point sources, scaled by the magnitude of annual emissions, and contours of the Douglas Pass nitrate factor source areas from Figure 13 (based on a spatial moving average analysis of nitrate factor concentrations associated with HYSPLIT back trajectory points).

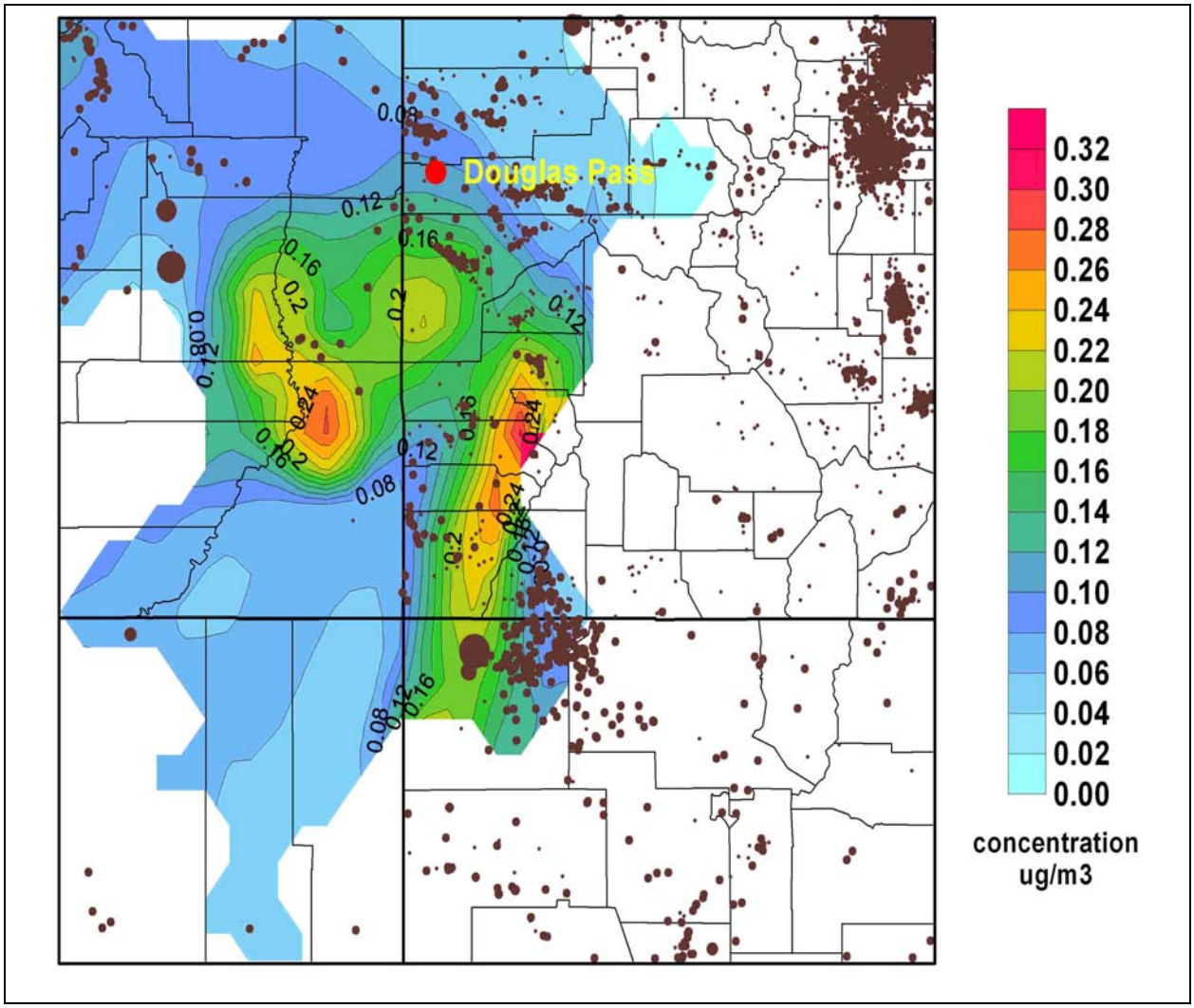

Figure 22. NOx point sources, scaled by the magnitude of annual emissions, and contours of the Douglas Pass nitrate factor source areas from Figure 13 (based on a spatial moving average analysis of nitrate factor concentrations associated with HYSPLIT back trajectory points).

Positive Matrix Factorization for Douglas Pass and Ripple Creek IMPROVE Data for 2004-2005

Methods and Model Diagnostics

Positive Matrix Factorization (PMF) is an analytical procedure that can be used to identify air pollution source categories and to estimate the contributions of these source categories for individual air quality samples in a time series. In this sense it is similar to Chemical Mass Balance analysis or CMB, but emissions or source profiles are not needed in PMF. PMF, which is similar to factor analysis, identifies source profiles through statistical analyses. Unlike factor analysis or principal component analysis, PMF yields non-negative estimates of the contributions of each source category to the total measurement for each sample. In the analysis reported here, PMF has been applied to IMPROVE aerosol fine mass concentrations.

The Environmental Protection Agency's (EPA's) PMF 1.1 software was used to estimate the contributions of four categories of sources for 2004 and 2005 for each of 220 sample-days at the Douglas Pass and Ripple Creek IMPROVE monitors. The EPA PMF 1.1 software requires non-missing, non-zero, and non-negative data for each sample and every aerosol species used in the analysis. EPA PMF 1.1 also requires uncertainty values for each species. Where possible, missing values and zero-value data were replaced with a number equivalent to half of the minimum detection level (MDL). Although there are other methods for estimating sample uncertainties for these data replacements (that may also yield better results), the reported IMPROVE sample uncertainties were used for all replaced data.

Hundreds of PMF 1.1 runs were completed for each of these two sites. This iterative process makes it possible to select a nearly optimal subset of analysis species and an optimal number of source categories for each site/data set. While this process yielded good results for the Douglas Pass data, the diagnostics for Ripple Creek showed that the results for that site were not useable. A brief summary of the results for Douglas Pass follows. A much briefer synopsis for Ripple Creek is included at the end of this section.

Twenty-two species were used in the final run for Douglas Pass. These species and the EPA PMF 1.1 signal to noise values for each are presented in Table 9. In order of appearance in Table 9, these fine mass species are aluminum, bromine, calcium, chromium, copper, elemental carbon, iron, total gravimetric fine mass, magnesium, nickel, nitrate, nonsoil potassium, organic carbon, lead, phosphorous, rubidium, selenium, silicon, sulfate, strontium, titanium, and zinc. Elemental carbon and organic carbon concentrations were calculated using the component carbon measurements and the IMPROVE formulas for these categories. See the IMPROVE Data Guide at:

http://vista.cira.colostate.edu/improve/Publications/OtherDocs/IMPROVEDataGuide/IMPROVEDataGuide.htm

Nonsoil potassium was also calculated using the formula presented in the IMPROVE Data Guide. Uncertainties for elemental carbon, organic carbon, and non-soil potassium were derived by taking the square root of the sum of the squares of the uncertainties for each component species used in the formulas for the concentrations.

Table 9. Species selected and PMF signal to noise values for the final PMF run for Douglas Pass.

Species	Signal To Noise
ALf	4.19
BRf	6.23
CAf	4.36
CRf	0.92
CUf	1.06
EC	1.1
FEf	8.69
MF	4.94
MGf	1.87
NIf	0.45
NO3f	7.76
NonsoilK	2.15
OC	1.64
PBf	1.18
Pf	0.53
RBf	0.63
SEf	0.74
SIf	4.33
SO4f	3.14
SRf	1.77
TIf	5.13
ZNf	3.19

EPA PMF 1.1 generates tables and plots for evaluating each run. Key diagnostics for the success or goodness of fit of the analysis are shown in Table 10. The values of Q (Robust and True) should be close to each other and not radically different from the product of the number of species times the number of samples. In this case, this product is 4840 (22 species times 220 sample days). The Q Robust and Q True numbers are close enough to 4840 to suggest a successful run. Q Robust and Q true are also very close to each other. The model run must also show convergence and Table 10 shows that it does.

Table 10 Key diagnostics for selected PMF run for Douglas Pass.

RANDOM RUN	Q (ROBUST)	Q (TRUE)	CONVERGED (Y/N)	STEPS
1	3501.2	3603.21	Yes	709

Additional diagnostics for the selected PMF run for Douglas Pass are presented in Table 11. Species with consistently low R-squared values in initial runs were assigned a "weak" species status in EPA PMF 1.1. EPA PMF 1.1 automatically triples the uncertainties for these species. Ideally, intercepts should be close to zero, slopes should be near 1, root mean square errors should be small compared to the species mean, and R-squared values should be near 1. Species with than less than ideal diagnostics can be included, but this may degrade the final resolution of the results. An additional data

uncertainty factor of 5% was selected for each use of PMF 1.1.

Table 11. Additional diagnostics for selected PMF run for Douglas Pass.

DIAGNOSTICS FOR POSITIVE MATRIX FACTORIZATION FOR DOUGLAS PASS IMPROVE DATA 2004-2005 (BEST RUN)				
SPECIES	INTERCEPT	SLOPE	ROOT MEAN SQUARE ERROR	R-SQUARED
Strong Species Used in Analysis:				
NO3f	0.01	0.94	0.02	1.00
FEf	0	1.00	0	0.99
SIf	-0.01	1.08	0.02	0.98
SO4f	0.05	0.91	0.06	0.98
CAf	0	0.9	0.01	0.96
TIf	0	0.98	0	0.96
BRf	0	0.85	0	0.95
MF	0.02	0.97	0.62	0.92
OC	0.04	0.83	0.21	0.88
ALf	0	0.85	0.02	0.85
MGf	0	0.71	0	0.83
EC	0.02	0.69	0.06	0.70
Pf	0	0.67	0	0.69
SRf	0	0.54	0	0.67
PBf	0	0.49	0	0.62
Weak Species Used in Analysis:				
CUf	0	0.38	0	0.55
RBf	0	0.45	0	0.52
NIf	0	0.52	0	0.46
Nonsoil Potassium	0	0.22	0.01	0.35
ZNf	0	0.32	0	0.27
CRf	0	0.16	0	0.24
SEf	0	0.42	0	0.13

Four source factors or categories were identified for Douglas Pass (2004 and 2005 data). The four source categories are a sulfate group, geologic group, wood combustion group, and nitrate group. Figures 23 through 26 show the mass of the species and percent of mass per species for each factor and the time series of contribution scalars. The sample scalar times the factor's data set mean is the sample concentration. The sulfate source factor is high in sulfate and relatively high in selenium, phosphorous, lead, and elemental carbon. The geologic source factor is very high in soil and geologic material metals and cations including aluminum, calcium, iron, magnesium, lead, silicon, strontium, and titanium. The wood combustion source factor is high in bromine, organic carbon, elemental carbon, and nonsoil potassium, which is an indicator for wood combustion. The wood combustion source factor is also somewhat high in selenium, and this suggests that there may be less than ideal separation between the PMF source categories for sulfate and wood combustion. The nitrate source factor is very high in nitrate and relatively high in elemental carbon, lead, phosphorous and zinc.

The sulfate group probably represents power generation sources and probably also

includes mobile sources and oil and gas combustion. The nitrate group probably represents power generation sources, mobile sources, and fuel combustion in general. There does not seem to be enough of a signal in the Douglas Pass data to separate these two groups into smaller categories. Model performance declined significantly when more than four categories were included. Both groups represent secondary aerosols.

The sulfate group shows weak seasonality with lower values in the winter and a few major episodes. The geologic group is lowest in the winter and highest in the spring, and this is consistent with the annual patterns for blowing dust. The wood combustion group is lower in the winter and higher in summer and late fall. This is consistent with the seasonal patterns for wildfire and controlled burning smoke. The nitrate group is highest in the late winter and spring, and there are a few major episodes in late 2005.

Figure 27 through 30 present another set of diagnostic plots for the EPA PMF 1.1 analysis. Each figure contains box and whisker plots of the variability in the percentage of species apportioned to a source factor and the variability in the mass of species apportioned for that factor. According to the EPA PMF 1.1 Users Guide (see references below), when "plus signs are clustered and are close to the box, then there is good reproducibility. The black star that is overlaid is the value from the base case run. Ideally, the black star should lie within the box. If this is not the case, then the base case likely has some observations that are atypical and these observations are influencing the factorization."

The PMF process is iterative in two ways. Any given model run involves many, internal iterations, starting from a random "seed". Through these iterations the model attempts to converge on a reproducible solution. In addition the model itself must be implemented many times before the user is likely to find the most appropriate set of species and factors. Good reproducibility refers to the fact that the internal iterations will probably yield the same result given the same inputs with each re-use of the model. This reproducibility increases the likelihood of a good fit with reality. Figures 27 through 30 generally show good reproducibility and minimal influences from extreme events, especially for the key species for each source group. The geologic source group, however, seems to show the effects of atypical events. This makes sense since blowing dust events, which may occur a few times a year, probably have a strong impact on the results for this category.

The final diagnostic for the analysis is shown in Figure 31. This is a linear regression of the sum of the concentrations from all four source categories against the measured or gravimetric fine-mass concentrations from the IMPROVE sample. The slope is near one at 0.97. This shows that on average the PMF results reproduce the total measured concentrations from the samples. The intercept is near zero, and the R-squared is high at 0.92.

The PMF results for Ripple Creek were not used. Although most of the diagnostics were good, the regression of total PMF concentrations against measured concentrations was weak. The slope was nearly one at 0.996, but the intercept was relatively high at 1.2 ug/m3, and the R-squared value was quite low at 0.34. There was too much scatter in the plot, suggesting that the modeled contributions would often provide a poor representation of reality.

Select References on Positive Matrix factorization (PMF):

Eberly, Shelly, 2005. "EPA PMF 1.1 User's Guide," U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC 27711, 32 pp.

Kim, E., Hopke, P.K., Edgerton, E.S., 2004. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. Atmospheric Environment 38, 3349-3362.

Kim, E., Hopke, P.K., Edgerton, E.S., 2003. Source Identification of Atlanta Aerosol by Positive Matrix Factorization. Journal of Air & Waste Management Association 53, 731-739.

Kim, E., Hopke P.K., 2004. Source Apportionment of Fine Particles in Washington, DC, Utilizing Temperature-Resolved Carbon Fractions. Journal of Air & Waste Management Association 54, 773-785.

Kim, E., Hopke, P.K., Larson, T.V., Maykut, N., Lewtas, J., 2004. Factor Analysis of Seattle Fine Particles. Aerosol Science and Technology 38, 724-738.

Kim, E., Hopke, P.K., Paatero, P., Edgerton, E.S., 2003. Incorporation of parametric factors into multilinear receptor model studies of Atlanta aerosol. Atmospheric Environment 37, 5009-5021.

Lee, Eddie; Chan, Chak K.; Paatero, Pentti, 1999. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong**,** Atmospheric Environment. Vol. 33, no. 19, pp. 3201-3212. 1999.

Paatero P., 1997, Least squares formulation of robust non-negative factor analysis, Chemometrics and Intelligent Laboratory Systems, Volume 37, Number 1, May 1997, pp. 23-35(13).

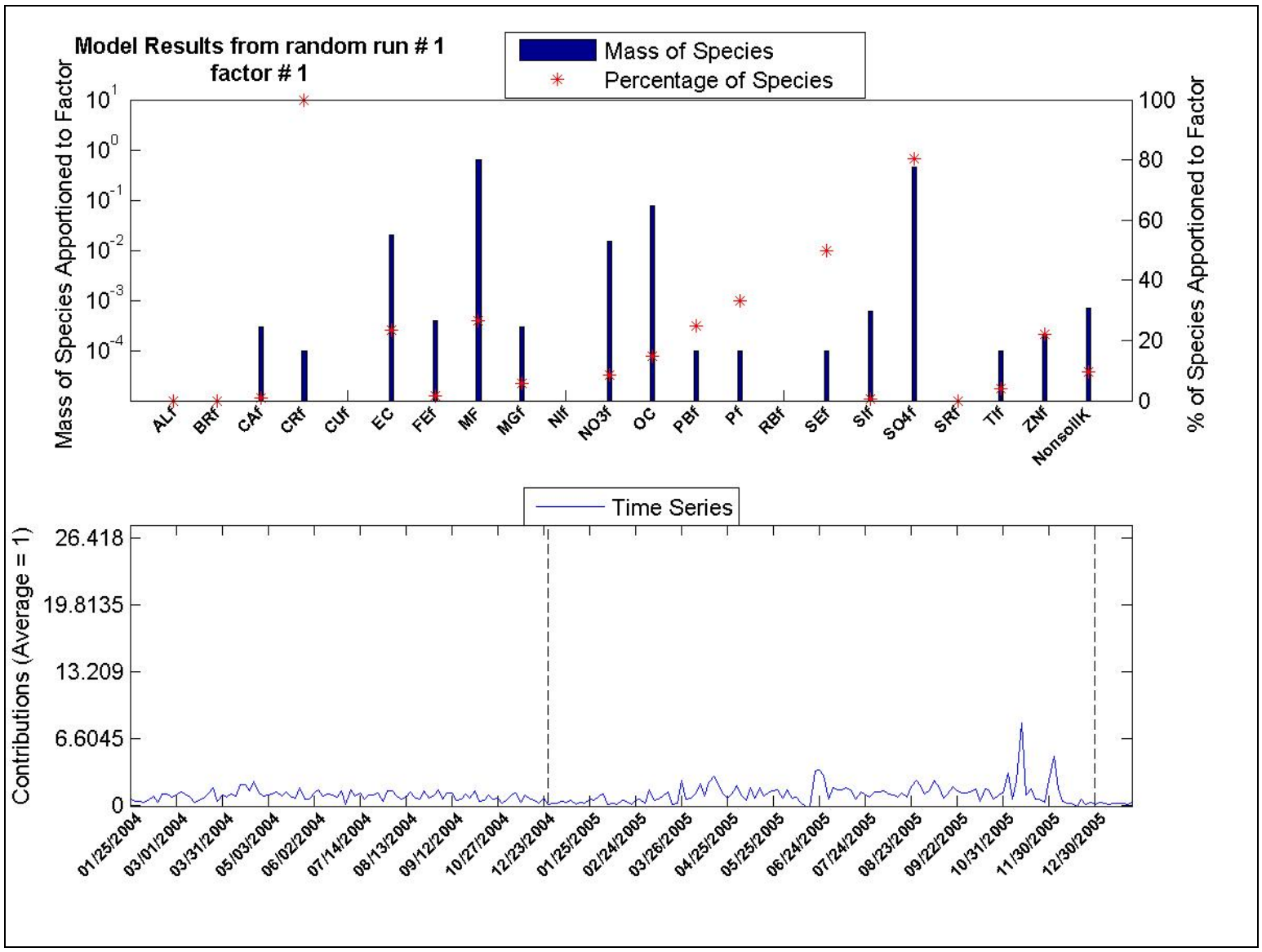

Figure 23. Mass of species and percent of mass per species for Factor 1 (Sulfate group) (top graph) and time series of contribution scalars (the sample scalar times the factor's data set mean is the sample concentration) (bottom graph).

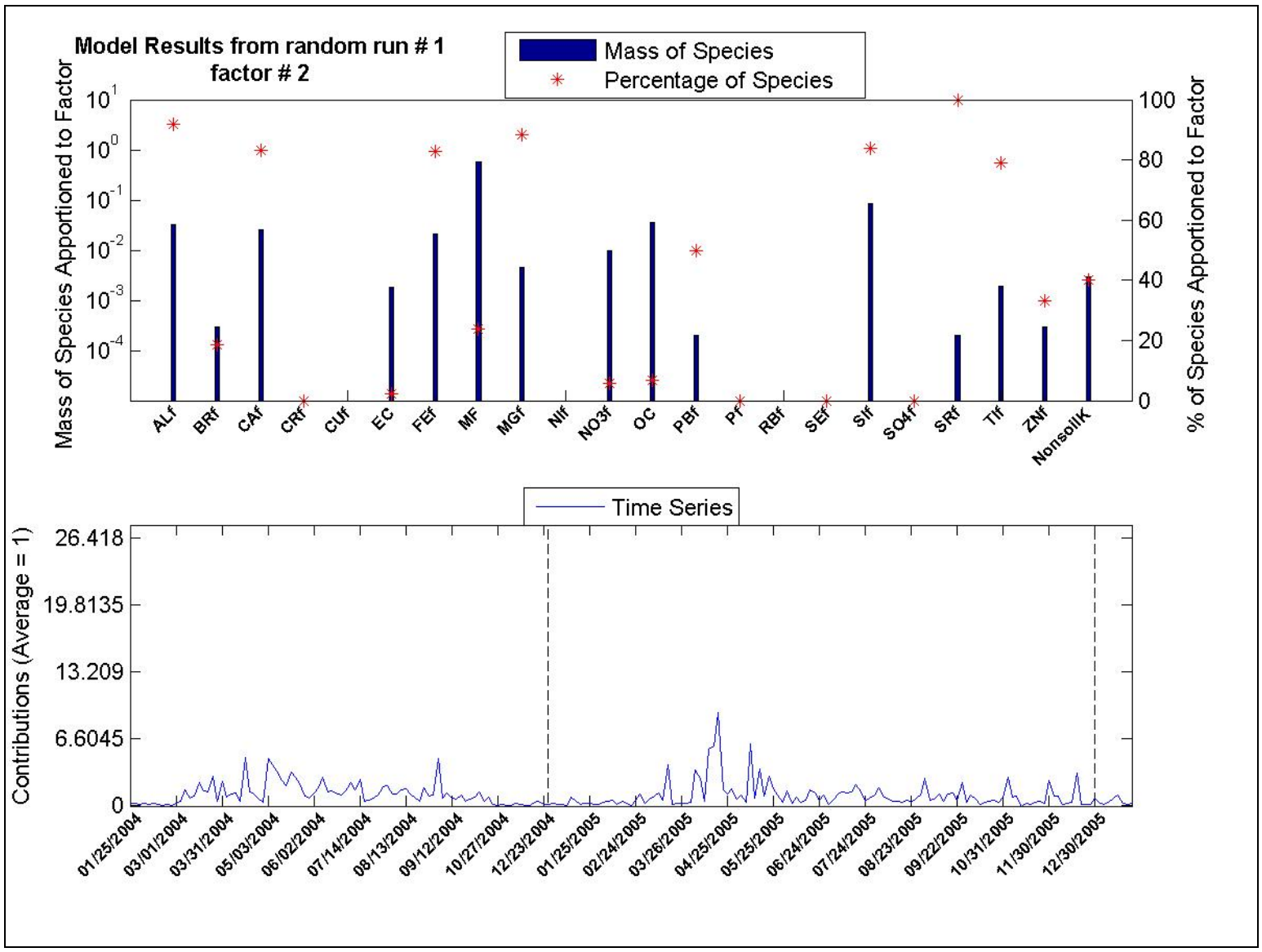

Figure 24. Mass of species and percent of mass per species for Factor 2 (Geologic group) (top graph) and time series of contribution scalars (the sample scalar times the factor's data set mean is the sample concentration) (bottom graph).

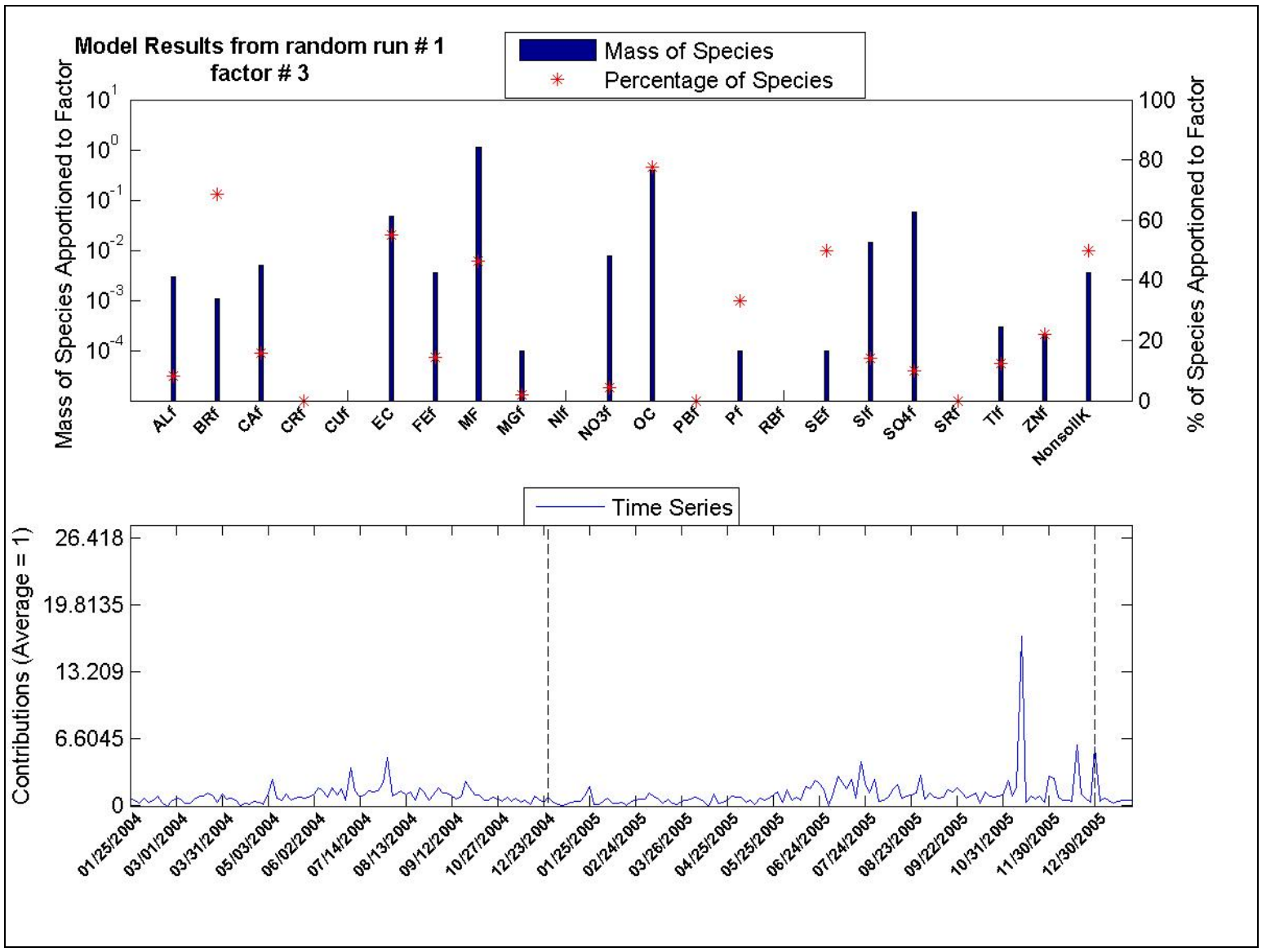

Figure 25. Mass of species and percent of mass per species for Factor 3 (Wood Combustion group) (top graph) and time series of contribution scalars (the sample scalar times the factor's data set mean is the sample concentration) (bottom graph).

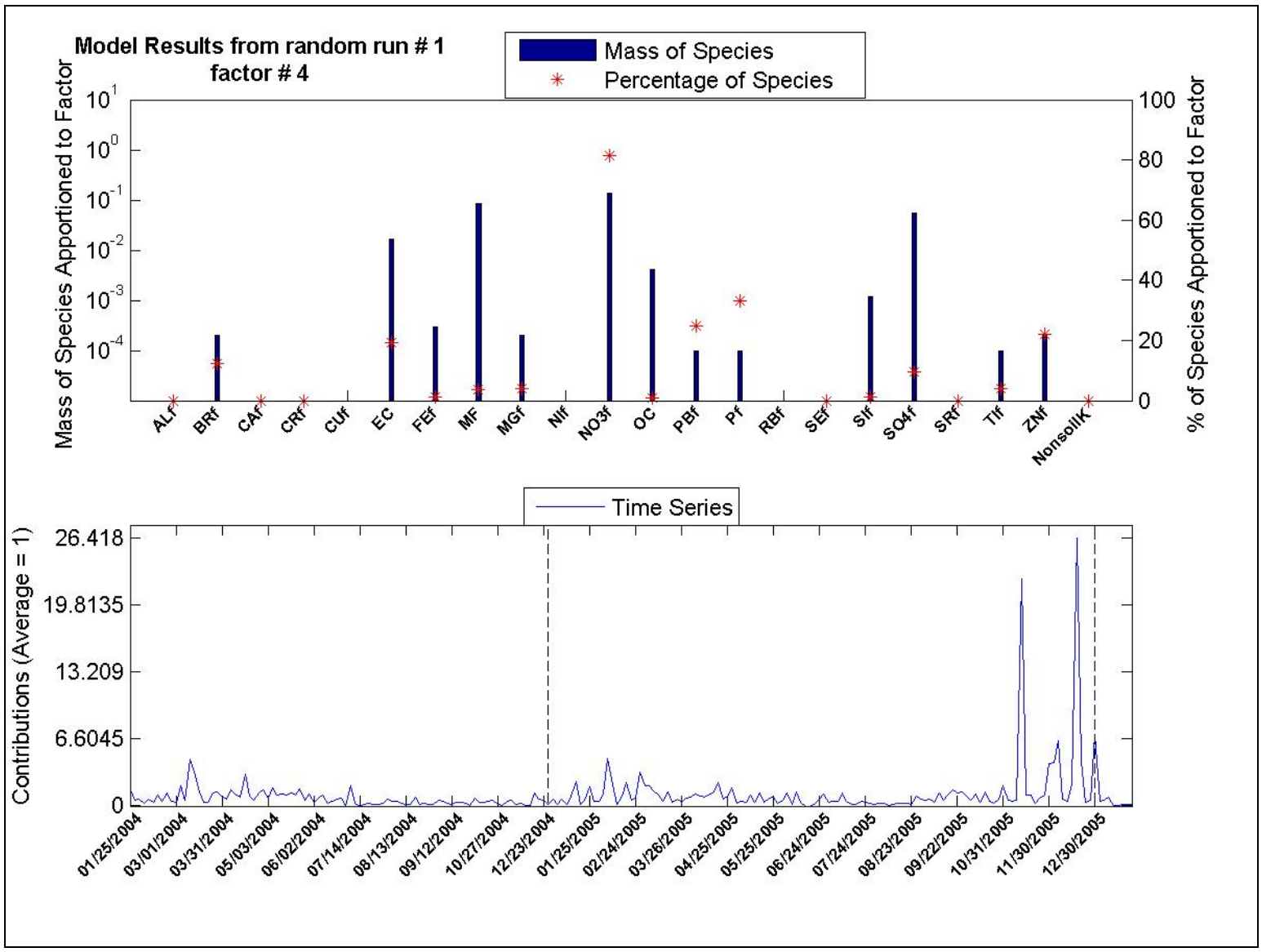

Figure 26. Mass of species and percent of mass per species for Factor 4 (Nitrate group) (top graph) and time series of contribution scalars (the sample scalar times the factor's data set mean is the sample concentration) (bottom graph).

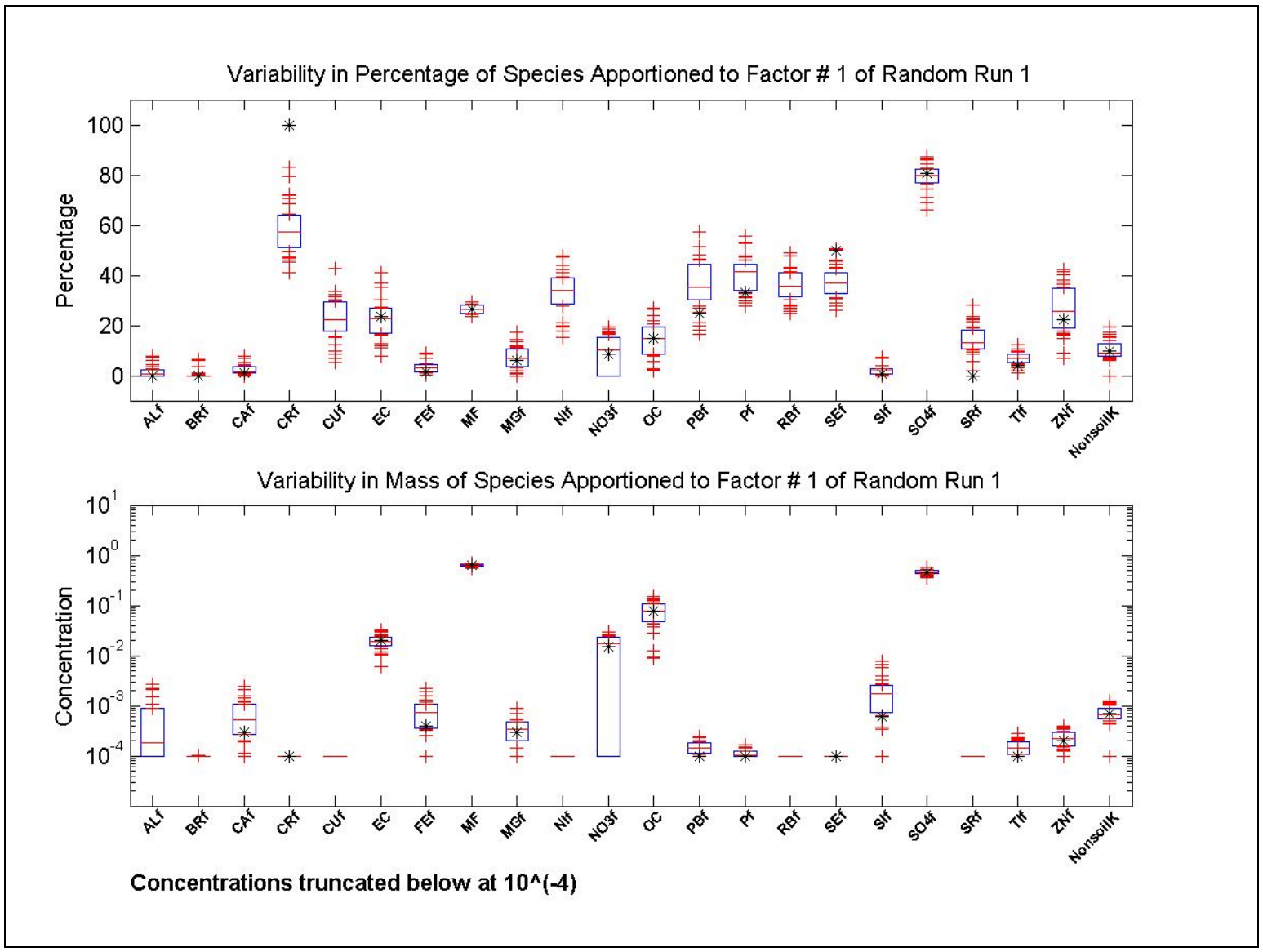

Figure 27. Box plots of variability in percentage of species apportioned to Factor 1, Sulfate group, (top) and variability in mass of species apportioned for Factor 1, Sulfate group, (bottom).

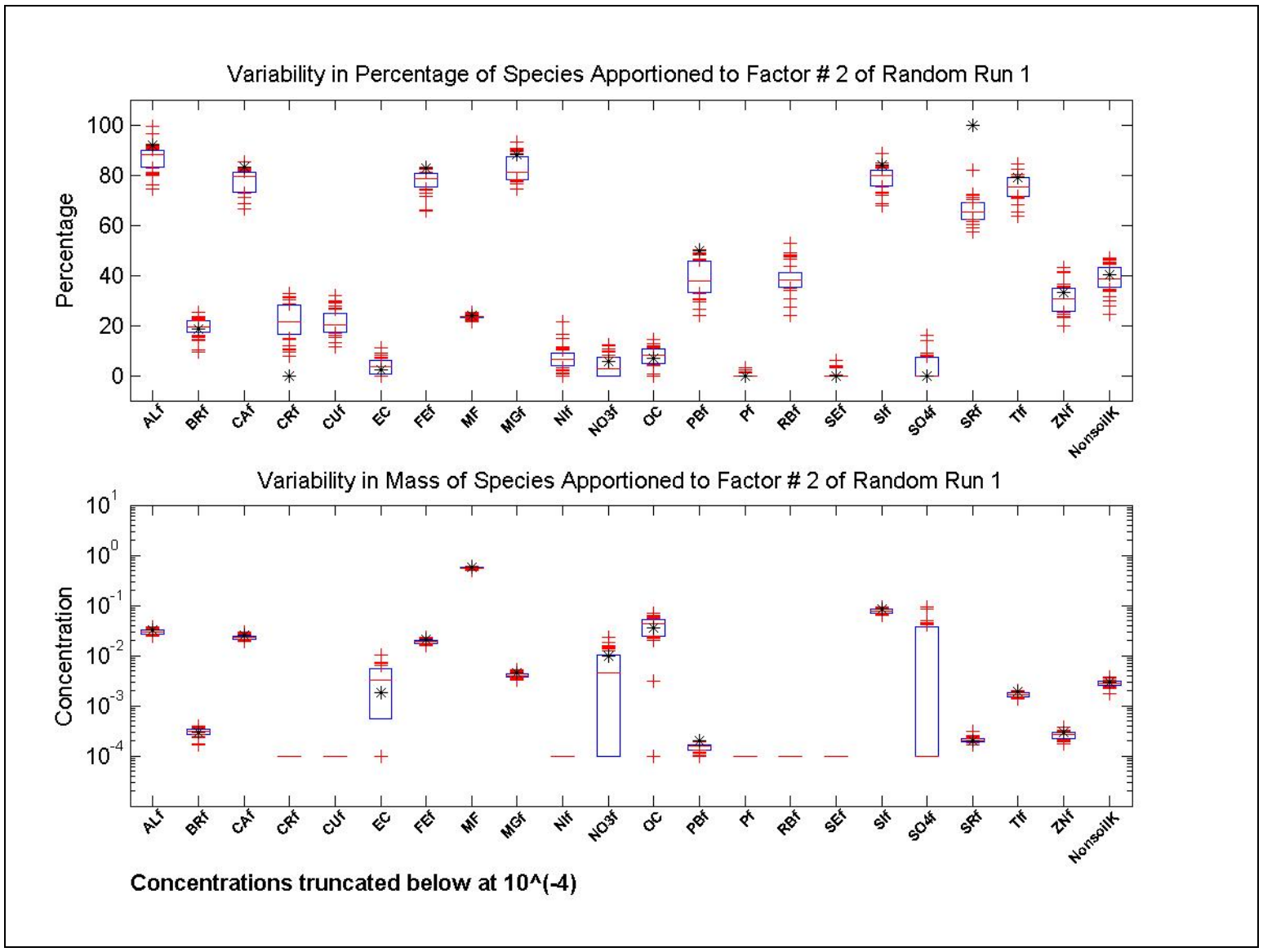

Figure 28. Box plots of variability in percentage of species apportioned to Factor 2, Geologic group, (top) and variability in mass of species apportioned for Factor 2, Geologic group, (bottom).

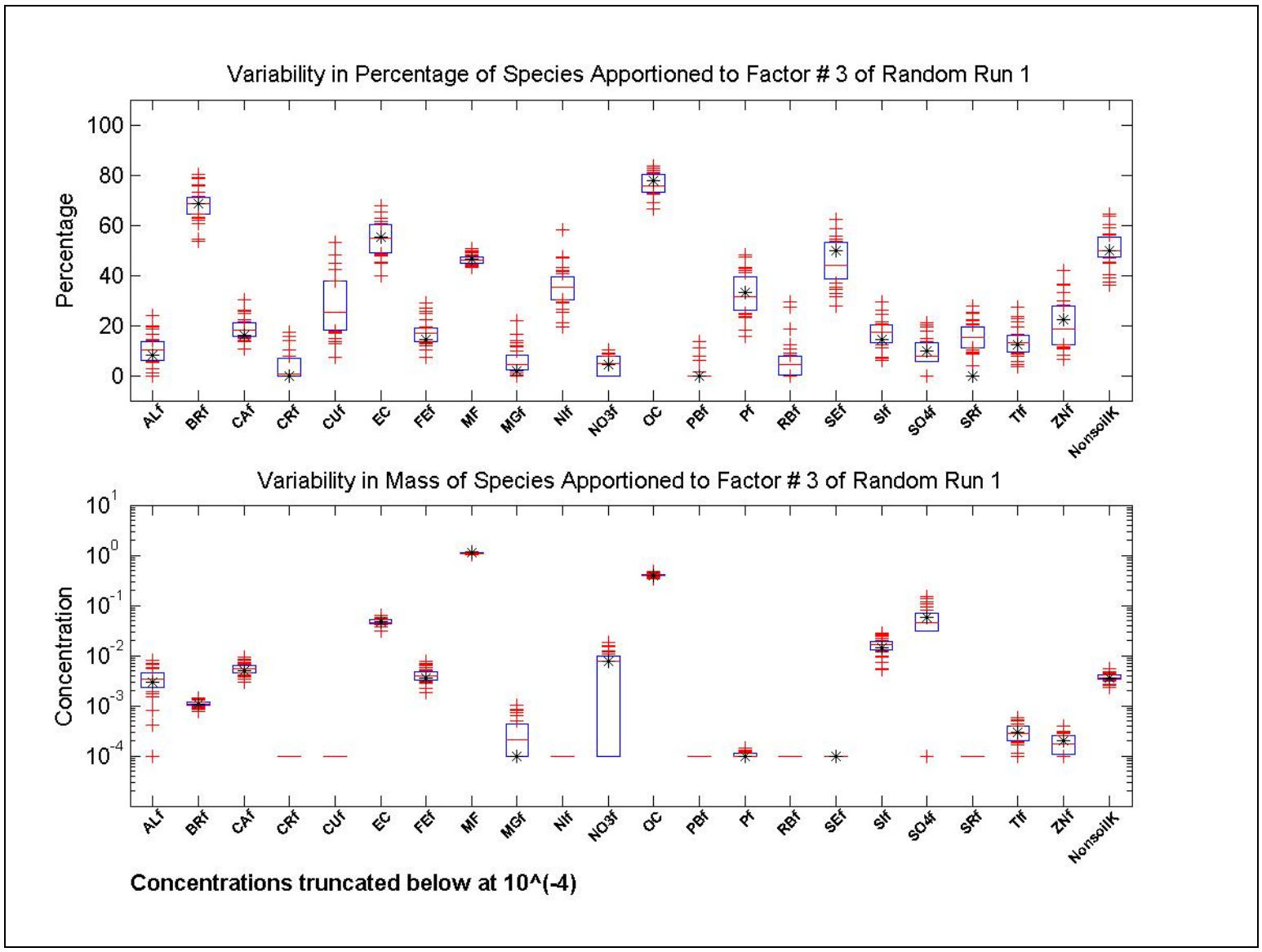

Figure 29. Box plots of variability in percentage of species apportioned to Factor 3, Wood Combustion group, (top) and variability in mass of species apportioned for Factor 3, Wood combustion group, (bottom).

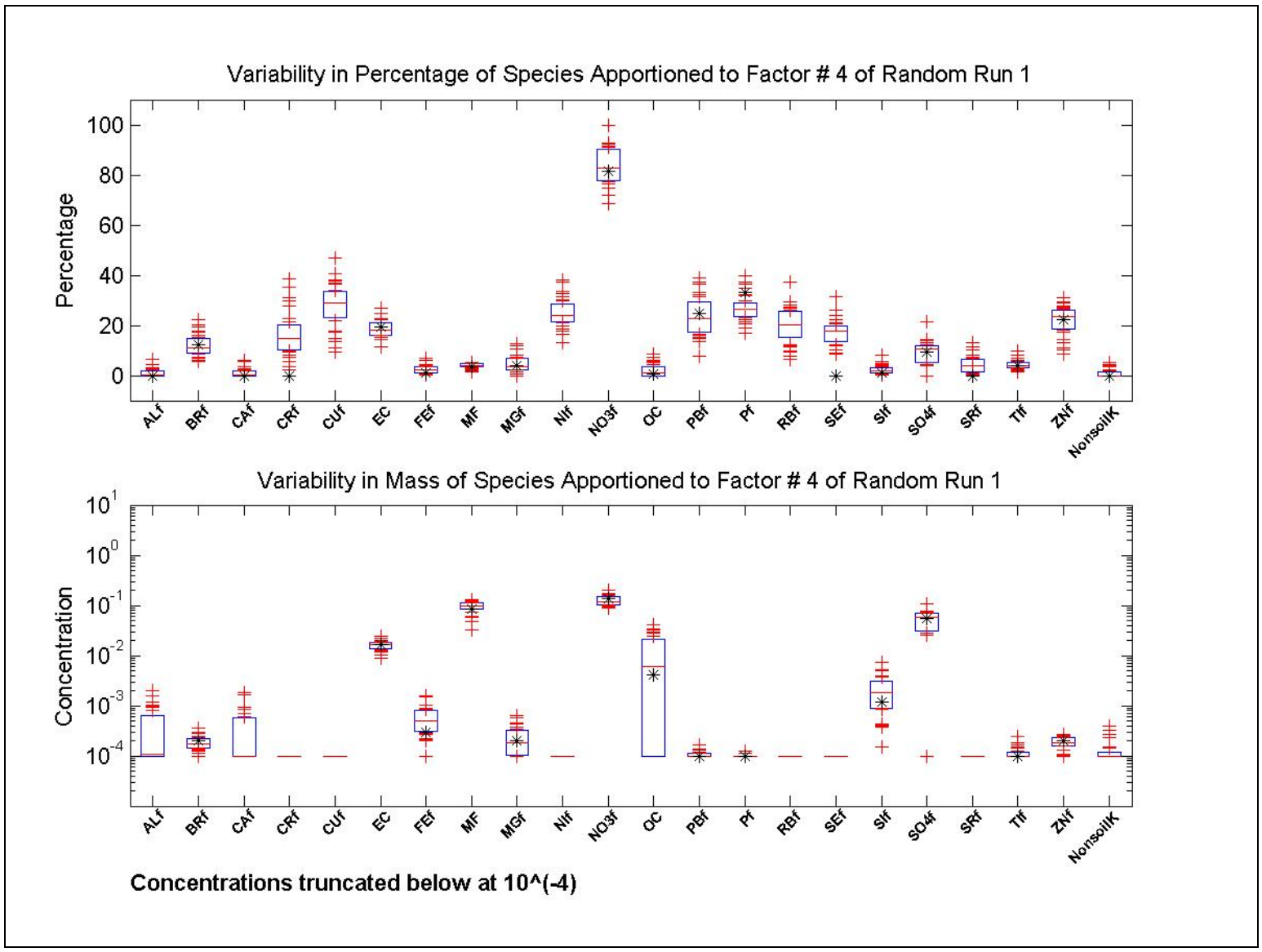

Figure 30. Box plots of variability in percentage of species apportioned to Factor 4, Nitrate group, (top) and variability in mass of species apportioned for Factor 4, Nitrate group, (bottom).

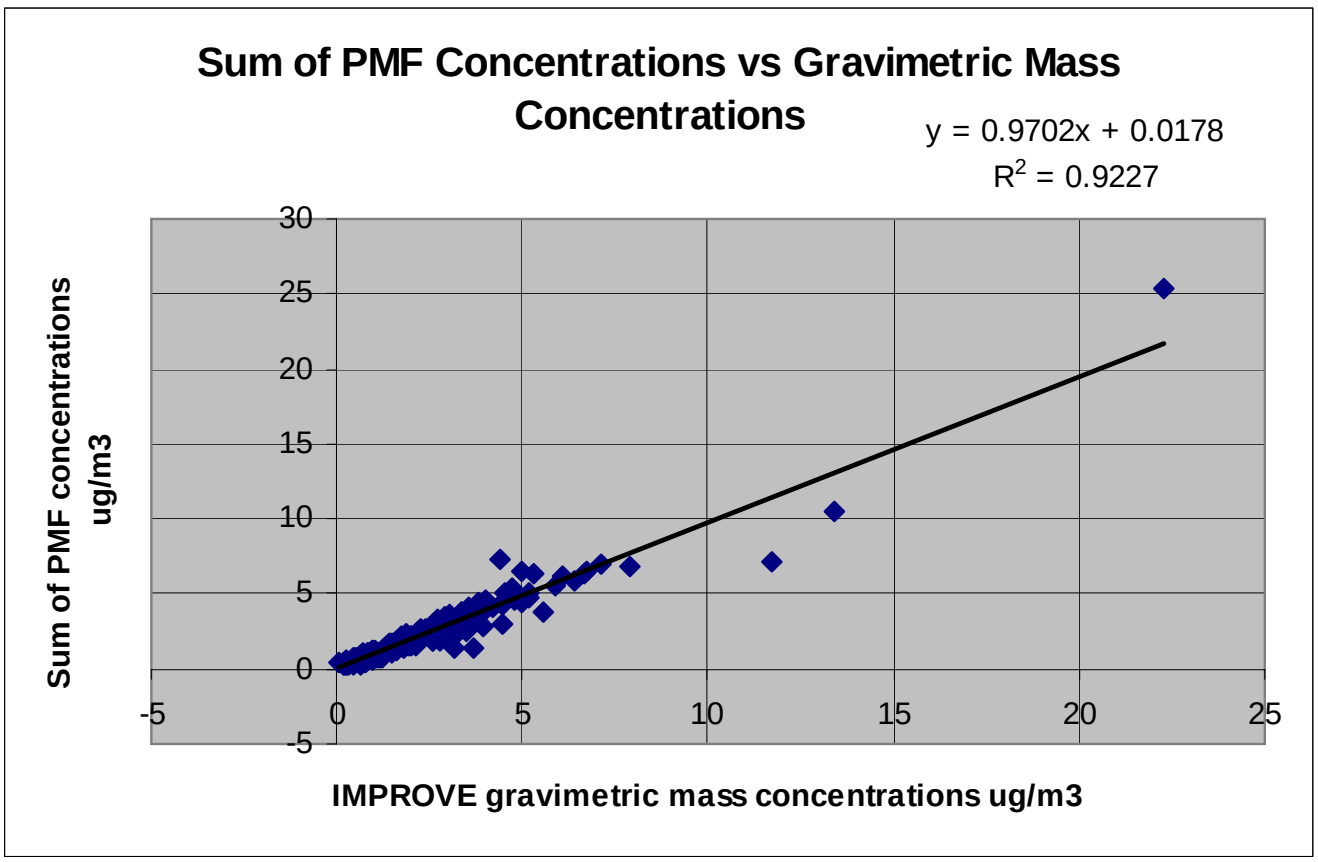

Figure 31. The sum of PMF concentrations versus gravimetric mass concentrations for Douglas Pass (2004-2005). The linear regression shows a nearly one-to-one correspondence and an R-squared of 0.92.