ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Atmospheric Environment ▮ (▮▮▮▮) ▮▮▮–▮▮▮

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US

W.D. Hafner*, N.N. Solorzano, D.A. Jaffe

University of Washington Bothell, Interdisciplinary Arts and Sciences, 18115 Campus Way NE, Bothell, WA 98011-8246, USA

Received 10 June 2006; received in revised form 29 September 2006; accepted 28 November 2006

Abstract

We calculated daily back-trajectories using the NOAA-HYSPLIT model to analyze 7 years of precipitation and PM2.5 data from three National Park sites in the Western US. Using a *k*-means clustering algorithm, the trajectories were segregated into six main transport patterns. At each site, we calculated trajectory clusters for 1, 5, and 10 days to represent short, medium and long-range flow patterns. Most clusters show marked seasonality. Faster flow patterns are more prevalent in winter, and slower/stagnant patterns are more prevalent in summer. The analyses between the 1, 5, and 10-day clusters revealed that the clusters of different duration show very different predictive power for rainfall and PM2.5. We found that the 1-day clusters are a better predictor for precipitation and PM2.5 concentrations, followed by the 5-day clusters. The 10-day clusters did a poorer job of differentiating precipitation and PM2.5. This is because the 10-day clusters show the greatest variability during the first day or two of transport.

Keywords: Aerosols; Cluster analysis; Trajectories

1. Introduction

The Western Airborne Contaminants Assessment Project (WACAP) was established in 2003 by the National Park Service in an effort to determine the degree and sources of contamination at alpine National Parks in the Western US. Although the project primarily exists for the measurement of semi-volatile organic contaminants and mercury, visibility problems caused by fine mass aerosols are also of concern to the National Park Service. Due to their inaccessible location, atmospheric transport and deposition are thought to be the primary pathway of contamination to these parks (Daly and Wania, 2005; Landers et al., 2003; Hageman et al., 2006). For this reason, modeling the transport of pollutants to the WACAP sites is an integral part of the project.

Although atmospheric pathways are highly variable, they tend to follow reproducible patterns as evidenced by any given location's climate. These pathways, or synoptic scale motions, tend to be active on scales of hundreds of kilometers. If individual synoptic motions can be determined, and grouped into patterns, these patterns can used

---

*Corresponding author. Tel.: +1 425 352 5357; fax: +1 425 352 5233.

E-mail addresses: whafner@uwb.edu (W.D. Hafner), djaffe@u.washington.edu (D.A. Jaffe).

1352-2310/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2006.11.049

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

to understand atmospheric transport to a given site, in conjunction with knowledge of emissions (Buchanan et al., 2002; Keim et al., 2005).

One method of identifying these patterns is to use cluster analysis on backwards air parcel trajectories. Trajectories are an approximation of the path that an air particle has taken to arrive at a particular site. In cluster analysis, trajectories that have similar transport speeds and directions are grouped together and assumed to represent specific synoptic patterns (Dorling and Davies, 1995; Kahl et al., 1997; Abdalmogith and Harrison, 2005; Mahura et al., 1999; Jaffe et al., 2005). Precipitation is often indicative of particular synoptic motions. By assigning precipitation amounts to the arrival time for each trajectory within each cluster, the assumption that clusters are representative of synoptic patterns can be evaluated. By assigning pollutant concentrations to the arrival time for trajectories within each cluster, an understanding can be gained about the most important routes of atmospheric contamination to a given site. (Sirois and Bottenheim, 1995; Brankov et al., 2003; Wang et al., 2004).

Although clustering has become a viable modeling technique, there is still some disagreement over choosing the ideal length of the back-trajectories used in the analysis. Many researchers choose the trajectory length based upon the atmospheric residence time of the of the pollutant being studied, whereas others choose a longer trajectory in order to more effectively model large scale meteorology (Kahl et al., 1997). To date, no one has studied whether clusters calculated with short or long trajectories are representative of the same patterns.

The goals of this paper are to apply cluster analysis to three WACAP sites. Precipitation and PM2.5 data will be applied to the cluster results in order to assist in answering these questions: Do the 1, 5, and 10-day trajectory clusters show similar transport, precipitation, and PM2.5 patterns? Which trajectory cluster (1, 5, or 10 day) gives the best representation of PM2.5 and precipitation data?

2. Methodology

2.1. Trajectory generation

The WACAP sites were selected to cover a wide latitudinal profile along the West Coast, as well as two parks from the interior (Landers et al., 2003). Of the eight primary sites, five have nearby measurements of PM2.5 and precipitation data available. This paper focuses on three of these sites: Rocky Mountain NP, Colorado; Mount Rainier NP, Washington; and Denali NP, Alaska. A full WACAP report, containing the analysis of these three sites plus Sequoia National Park, California and Glacier NP, Montana, will be available in fall 2007 on the WACAP website (http://www2.nature.nps.gov/air/Studies/air_toxics/wacap.cfm). Within each park, two alpine watersheds were chosen for the collection of WACAP samples. Trajectories were calculated for a location that was midway between the two watersheds. The coordinates and altitudes of these locations are presented in Table 1.

All back-trajectories were generated using the PC Version 4.7 of the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Hess, 2004) with the National Center for Environmental Prediction (NCEP) data grid. NCEP is a global 2.5° latitude/longitude grid with 18 vertical levels archived every 6 h. At each site, 3D kinematic back-trajectories were generated daily at 1800 UTC from 1998 through 2004 for durations of 1, 5, and 10 days. This is a total of 2557 trajectories per site per trajectory length. Each of the trajectory output files consists of latitude, longitude, and elevation coordinates every hour (hourly points) along the trajectory's path.

Table 1
Distance and direction of the precipitation (SNOTEL or CASTNET) and IMPROVE stations relative to the coordinates of the WACAP sites

Site	WACAP starting location			Precipitation station		IMPROVE station	
	Longitude	Latitude	Altitude (m)	Distance	Altitude (m)	Distance	Altitude (m)
DENA	−151.30	63.30	660	126 km E	660	125 km E	660
MORA	−121.90	46.85	1560	14 km SE	1560	20 km SW	430
ROMO	−105.64	40.29	3100	20 km NW	3260	8 km E	2760

All altitudes are above sea level.

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

There was some difficulty in selecting starting heights for the back trajectories. The HYSPLIT PC model only allows for calculation of trajectories above the model ground level (AGL). Given the NCEP grid size, there could be substantial differences between the real altitude of the WACAP sites and the NCEP model ground level. Taking advice from the HYSPLIT web site, we used the difference between the real-world site elevation and the model ground level as the trajectory starting height. With the exception of Denali in Alaska, the sites are located in mountains (see altitudes in Table 1) resulting in trajectory starting heights that are higher than the NCEP model ground level. The WACAP sites at Denali are located in glacial flats a few kilometers north of the Alaska Range. This location is below the NCEP model ground level. At Denali, 0 m AGL was used as the starting height.

2.2. Cluster analysis

The goal of cluster analysis is to group trajectories so as to minimize the variability of trajectories within a cluster, and maximize the variability between clusters. The method chosen for this analysis was a non-hierarchical *k*-means algorithm that had previously been used to determine the long-range climatology to the Azores Islands (Owen, 2003).

One aspect of using non-hierarchical clustering is that the number of clusters must be pre-determined. Although there are methods to aid in picking the "correct" number of clusters (Dorling et al., 1992), the process remains subjective. Since it was our goal to compare transport patterns between trajectory lengths and between sites, we wanted to use a consistent number of clusters. The selection criteria were to have enough clusters to represent the different primary transport patterns, yet not so many as to obscure the analysis. The number of clusters commonly used in the literature ranged from 5 to 11. We tried using 5, 6, and 7 clusters per site, and after a visual inspection, settled on 6 as the best representation of averaged meteorological patterns.

There are four steps to the *k*-means algorithm. The first involves creating initial partitions, or seeds, amongst the trajectories. These initial seeds are created by randomly selecting six trajectories from the total. In the second step, distance (in degrees) between the hourly points of the individual trajectories and the seed trajectories is calculated. Trajectories are assigned to the seed with which they share the smallest sum distance (the seed they are closest to). In the third step, a cluster center is calculated for each of the six groups of trajectories. This creates a new set of slightly more accurate seed trajectories. The fourth step is to continually repeat steps 2 and 3 until no trajectories change cluster assignments. In this manner, the *k*-means method allows for the trajectories to change cluster membership until the best fit is obtained (Owen, 2003). Once the trajectories are sorted, the standard deviation for each hour is calculated between the member trajectory's location and cluster mean. In this manner, a measure of both horizontal and vertical variance is determined about the cluster mean.

An important aspect of clustering is the effect of altitude. Running the cluster model with and without including altitude in the distance calculation seems to have little direct impact in assigning the trajectories to their clusters (Owen, 2003). However, altitude is the primary factor in trajectory length, as wind speeds tend to be greater at higher altitudes (Harris et al., 2005). If there were several trajectories from the same direction but of different lengths, the result would be separate clusters forming for the higher, fast moving trajectories and the slower, lower altitude trajectories.

2.3. Precipitation and aerosol data

After calculating the clusters, we assigned daily local precipitation and every third day PM2.5 data to the clusters associated with that day's trajectory. Daily precipitation data were collected by the Snopack Telemetry Network (SNOTEL; http://www.wcc.nrcs.usda.gov/snow/). A SNOTEL station was not present near Denali in Alaska. Instead, precipitation data from a nearby Clean Air Status and Trends Network (CASTNET) station was used. To ensure that the precipitation data was representative of our sites, the chosen stations are all at an altitude similar to that of the WACAP sites (see Table 1). SNOTEL data is archived as daily measurements ending at midnight. Only 6 days of SNOTEL data were missing for the whole 7-year period at each of the sites, resulting in nearly complete data sets. For Denali, hourly CASTNET data was converted to daily totals for this analysis. A total of 2828 h were missing, with the bulk of these occurring from February through April 2003, and September 2004.

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

Each of the National Parks in this study also is home to an Interagency Monitoring of Protected Visual Environments (IMPROVE) sampling station (http://vista.cira.colostate.edu/improve/Default.htm). At each site, aerosols are collected on filters in two different size ranges: less than 2.5 μm and between 2.5 and 10 μm aerodynamic diameter. This paper is only concerned with the fine particulate matter, or PM2.5. These 24 h samples were collected from midnight to midnight local time. Originally, the schedule was such that samples were collected 2 days per week, on Wednesday and Saturday. Beginning late in the summer of 2000, this schedule changed so that samples were collected every third day. Table 1 gives the altitude for each of the IMPROVE stations, as well as the distance and direction from the WACAP sites. The IMPROVE station at Mount Rainier is at a significantly lower elevation than the WACAP site. While there are undoubtedly differences in transport between these altitudes, they are small compared to the larger uncertainties of the NCEP grid size.

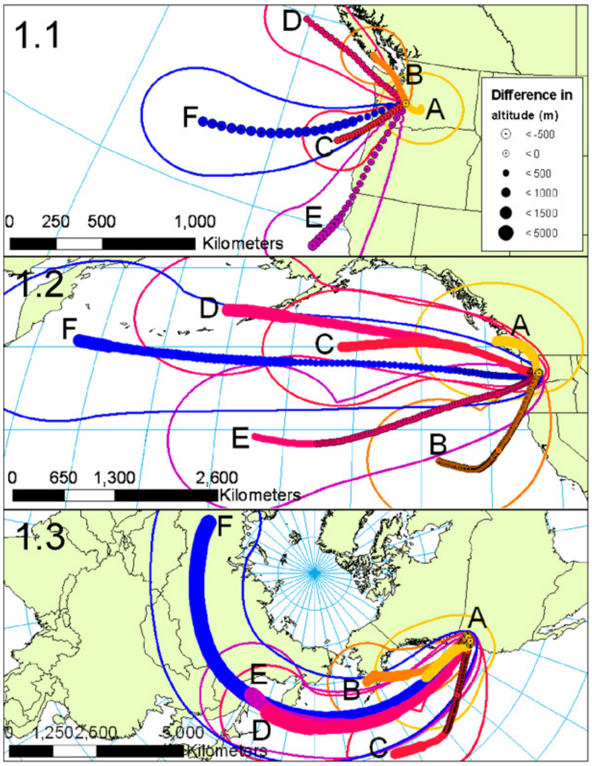

Fig. 1. Cluster centers for 1 (1.1), 5 (1.2) and 10-day (1.3) back trajectory clusters at Mount Rainier National Park. The cluster centers are color coded by length and labeled alphabetically, with A being the shortest and light orange, B being darker orange, C is red, D is light purple, E is purple, and F being the longest and blue. Horizontal standard deviations are presented for each cluster center in the same color as the cluster. Altitude is represented by different sized points. Color versions of all figures are available online.

3. Results and discussion

3.1. Seasonality and precipitation

The result of our clustering analysis was nine separate sets of clusters (three sites by three trajectory lengths), with six individual clusters in each set. Each of these clusters contained between 58 and 1047 trajectories. The number of trajectories per cluster divided by 2557 (7 years × 365 days) is referred to as the trajectory membership. Figs. 1–3 show these clusters as represented by their cluster centers; or the average position of all the trajectories within a particular cluster. The cluster centers are organized alphabetically and color-coded by length with A being the shortest, F the longest (color versions of all figures are available online through Science Direct). The relative altitudes of the hourly points in the cluster mean are also included in these figures. These were calculated by subtracting the trajectory starting height from the hourly point altitude.

It must be remembered that although these cluster centers appear to show one atmospheric pathway, the trajectories within each cluster can exhibit substantial variability about the mean. All figures illustrate this point by showing the horizontal standard deviation about each cluster center, forming what can better be thought of as an envelope about the mean (Sirois and Bottenheim, 1995; Kahl et al., 1997). In general, vertical standard deviations for the 1-day clusters average 100–300 m; 5-day clusters average 400–1500 m; 10-day clusters 500–1800 m. The vertical standard deviations are lower near the trajectory start and higher near the end points.

In Table 3, we compare the fraction of the total precipitation in each cluster (cluster total divided by 7 year total) to the trajectory membership. For example, 1-day cluster F at Mount Rainier has a trajectory membership of 6.7% but is responsible for 25.1% of the precipitation at the site (see Table 3). This enhancement of precipitation implies that the individual trajectories of this cluster are influenced by similar synoptic motions. The seasonality of the clusters, as shown in Table 2, is defined as December, January, and February

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

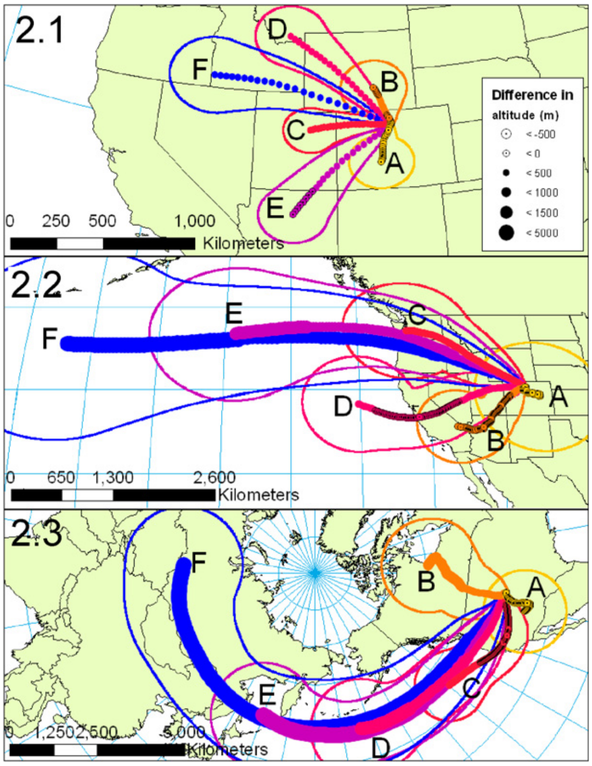

Fig. 2. Cluster centers for 1 (2.1), 5 (2.2), and 10-day (2.3) back trajectory clusters at Rocky Mountain National Park. Colors are the same as Fig. 1. Horizontal standard deviations are presented for each cluster center in the same color as the cluster. Altitude is represented by different sized points.

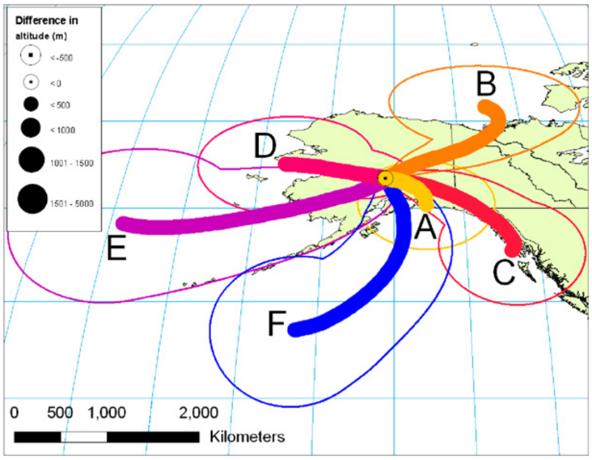

Fig. 3. Cluster centers for the 5-day trajectories at Denali National Park. Colors are the same as Fig. 1. Horizontal standard deviations are presented for each cluster center in the same color as the cluster. Altitude is represented by different sized points.

representing winter, and the rest of the seasons defined accordingly.

Each cluster represents trajectories from different seasons over a 7-year period giving rise to the question: Does every day in a particular cluster represent similar synoptic conditions? The aforementioned cluster F at Mount Rainier contains 86 winter trajectories and only five summer trajectories (see Table 2.1). For these 5 summer days, the averaged 850 mb geopotential height at Mount Rainier is 1430 m. Compared to the long-term summer average of 1510 m, it is apparent these 5 summer days are not typical, and have more in common with the predominantly winter trajectories of this cluster. Fig. 4 shows the 850 mb geopotential height plot for one of these days, 12 June 2000, compared with the 850 mb plot for the summer average and winter average.

Mount Rainier is located in the heart of the Pacific Northwest, a region that sits astride the midlatitude cyclonic storm belt (Jackson, 1993). A visual inspection of the 1 and 5-day cluster centers in Figs. 1.1 and 1.2 shows recognizable patterns. Cluster centers A for the 1 and 5-day clusters are representative of easterly flow to the site, although it must be noted that the standard deviation indicates that many of these trajectories circulate around Mount Rainier (see Figs. 1.1 and 1.2). While cluster A in both the 1 and 5-day clusters contains some trajectories from all seasons, they are mainly composed of trajectories from the summer (see Table 2.1). 1-day cluster B and 5-day cluster C have an even stronger summer presence than cluster A.

As expected from the offshore presence of the Pacific High in the summer, July and August are the driest months at Mount Rainier. The short 1-day clusters A and B have a combined trajectory membership of nearly half (48%) of the trajectories, but only contribute 10% of the precipitation to Mount Rainier. Similarly, the 5-day clusters A and C have over 45% of the trajectories, but only 23% of the precipitation (see Table 3).

Cluster centers E and F of the 1-day plots depict cyclonic flow from the southwest, a trend that is also seen in clusters B and E of the 5-day plots (see Figs. 1.1 and 1.2). Table 2.1 indicates that all four of these clusters are comprised of trajectories from the winter months, with only a minimal contribution from the summer. This is especially true of the 1-day clusters E and F, which have 159 and 86 winter trajectories and only 2 and 5 summer trajectories, respectively.

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

Table 2
Seasonal trajectory totals from each cluster for (2.1) Mount Rainier and (2.2) Rocky Mountain National Parks

	2.1 Mount Rainier National Park							2.2 Rocky Mountain National Park					
	A	B	C	D	E	F		A	B	C	D	E	F
1-Day													
WI	101	72	149	65	159	86	WI	29	70	148	149	94	142
SP	140	140	188	89	44	43	SP	87	119	161	96	116	65
SU	179	294	107	57	2	5	SU	238	137	191	18	52	8
AU	172	129	147	95	57	37	AU	113	72	205	93	100	54
5-Day													
WI	62	153	101	84	128	104	WI	18	147	126	144	127	70
SP	142	92	146	122	100	42	SP	79	173	147	132	83	30
SU	197	29	250	87	66	15	SU	175	299	105	49	16	
AU	124	72	150	130	80	81	AU	81	233	148	64	73	38
10-Day													
WI	98	78	116	130	153	57	WI	74	10	199	182	128	39
SP	139	108	116	136	120	25	SP	70	45	286	159	68	16
SU	170	123	150	138	59	4	SU	231	25	322	59	7	
AU	121	110	90	131	133	52	AU	180	13	240	114	62	28

From Table 3, we see that there are large differences at Mount Rainier between the trajectory memberships and precipitation amounts for these predominantly winter clusters. From 1998 to 2004, 1960 cm of precipitation fell at Mount Rainier. For the 1-day clusters, over 50% of this precipitation is associated with clusters E and F, even though they contain 17% of the total trajectories. With the 5-day clusters, B and E have nearly 50% of the precipitation, but only 28% of the trajectories.

The same trends between trajectory membership and precipitation that are seen for the 1 and 5-day clusters are not seen for the 10-day clusters (see Fig. 1.3). For all of the 10-day clusters at Mount Rainier, the difference between the trajectory membership and relative precipitation are negligible (see Table 3). Some seasonality does exist within the 10-day clusters; particularly with clusters E and F, which represent the fast moving and higher altitude air masses of the winter months (see Table 2.1). However, seasonality is not as consistent as it was with the 1 and 5-day cluster lengths. It seems as though the longer trajectories represent atmospheric motion on a planetary scale. The clustering algorithm sees these longer, straighter, trajectories and clusters them accordingly.

The variance about the cluster mean is much higher in the first 100–500 km of the 10-day clusters than it is for the 5-day clusters, which in turn is higher than the 1-day clusters. This high variance within a few hundred kilometers coincides with the same scale as synoptic scale motion. A reduced ability to group trajectories on synoptic scale motion is a likely reason 10-day clusters at Mount Rainier are not as representative of precipitation. Although not presented in this paper, Glacier and Sequoia National Parks displayed a similar trend of declining precipitation variability between clusters made with longer duration trajectories.

Colorado in general, and Rocky Mountain National Park specifically, are located distant from the Pacific Ocean and the Gulf of Mexico, the continent's major sources of precipitation. In addition to its continental location, Colorado is the highest contiguous state in the US, containing over $\frac{3}{4}$ of the nation's land above 3000 m (Doesken et al., 2003). The distance minimizes the potential synoptic effects from either precipitation source, and the high topographic variability increases the frequency of locally generated storms, or convective precipitation (Baron and Denning, 1993; Doesken et al., 2003). A result of these factors is a similar rainfall for all seasons. The 7 year totals at Rocky Mountain are 170 cm for winter, 178 cm for spring, 123 cm for summer and 117 cm for fall. At both Mount Rainier and Denali, the total precipitation for the wettest season was at least a factor of 5 greater than the driest season, contributing to the differences in precipitation between clusters at these sites.

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

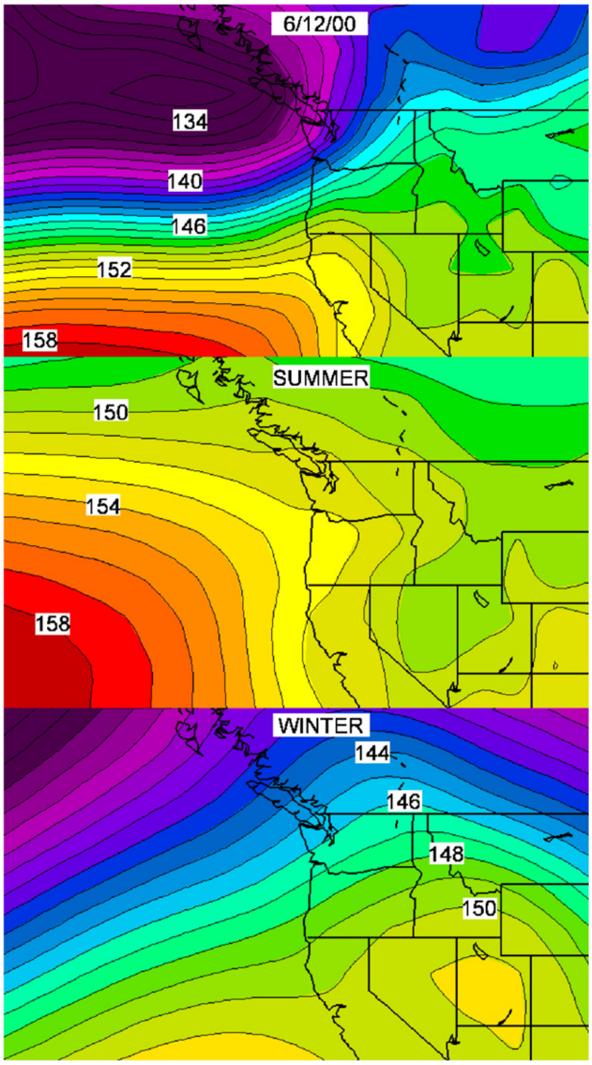

Fig. 4. 850 mb geopotential heights for 12 June 2000, a member of 1-day cluster F at Mount Rainier, and the long-term winter and summer climatological averages.

The trend of similar seasonal rainfall at Rocky Mountain means there is little difference between the trajectory membership and the percent of precipitation for each cluster. This is true for the 1, 5, and 10-day trajectory clusters. Although precipitation trends do not exist, seasonality (see Table 2.2) is present in the 1, 5, and 10-day cluster centers (see Figs. 2.1, 2.2, and 2.3, respectively). The shortest clusters, A and B, are dominated by days from the summer months, and the longest clusters, E and F, are primarily composed of trajectories from winter (see Table 2.2).

Clusters A and B for the 1-day plots and cluster A from the 5 and 10-day plots have some precipitation and contain more summer trajectories than other seasons (see Table 2.2). Although it is tempting to suggest that these clusters are representative of convective precipitation, this is not likely given the NCEP grid size.

Fig. 3 shows the 5-day cluster centers for Denali National Park. The Denali clusters are unique in that they have the shortest average cluster length of all sites for the 1, 5, and 10-day cluster centers. In addition, the 5 and 10-day cluster centers at Denali contain southerly and easterly flows, as opposed to predominately westerly flows seen at the other sites. Harris and Kahl (1994) conducted a similar clustering analysis for Barrow, Alaska using 10-day back trajectories and trajectory starting heights of 500, 1500, and 3000 m above sea level. Their results were similar to what we found: their cluster centers were typically short, with a portion of flow from non-westerly directions. This is in part due to the high latitude site, which is frequently is under the influence of the polar easterlies, which occur pole ward of the Aleutian low pressure system (Harris and Kahl, 1994). However, an additional cause for the range in trajectory directions is the starting height used at Denali. Denali has the lowest altitude of the WACAP sites, and was the only site where the trajectories were started at NCEP model terrain height, or 0 m above ground level (see Table 1). In their comparison of trajectory starting heights, Harris and Kahl (1994) found that trajectory clusters started at 500 m contain easterly flow, which diminishes for the 1500 m start height, and disappears with the 3000 m start height.

3.2. Analysis of PM2.5 data with trajectory clustering

We also evaluated the IMPROVE PM2.5 data with the clustered trajectories. First, we first log-transformed the data to improve the normality of the distribution. Next, we calculated the average PM2.5 concentrations for each site and each cluster. We then used a *t*-test to determine whether the average cluster concentrations were significantly different than the average for each site. All average concentrations were calculated for 1998–2004. To ease interpretation of the results, the concentrations and standard deviations presented in Table 3 and in the text are the anti-logs of the averages. Bold values in Table 3 denote clusters that are significantly

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

Table 3
Trajectory membership, relative amount of precipitation, and the geometric mean PM2.5 concentration and standard deviation per cluster in $\mu g\,m^{-3}$

		A	B	C	D	E	F
1-day clusters							
Memb.	Denali	19.4%	13.1%	18.9%	23.7%	15.8%	9.2%
Precip.	31.4	9.0%	36.7%	40.1%	3.3%	7.5%	3.5%
PM2.5	0.9±2.6	**1.2±2.2**	**1.2±2.6**	**1.3±2.6**	**0.6±2.5**	**0.7±2.4**	**0.6±2.6**
Memb.	Rainier	23.2%	24.8%	23.1%	12.0%	10.2%	6.7%
Precip.	280.0	4.9%	6.1%	27.5%	11.4%	25.1%	25.1%
PM2.5	2.8±2.4	**4.9±1.9**	**4.5±1.7**	**2.2±2.1**	2.6±1.7	**1.1±2.3**	**1.0±2.3**
Memb.	Rocky	18.3%	15.6%	27.6%	13.9%	14.2%	10.5%
Precip.	84.3	19.8%	17.4%	23.0%	13.2%	13.1%	13.4%
PM2.5	2.2±2.5	**3.8±1.9**	**2.7±2.2**	2.3±2.4	**1.4±2.2**	2.2±2.4	**1.0±2.5**
5-day clusters							
Memb.	Denali	34.6%	12.9%	17.4%	15.5%	6.4%	13.2%
Precip.	31.4	25.5%	2.4%	8.6%	19.5%	22.7%	21.4%
PM2.5	0.9±2.6	1.0±2.4	1.1±2.3	**0.5±2.4**	**1.4±2.5**	1.0±2.9	0.8±2.6
Memb.	Rainier	20.5%	13.5%	25.3%	16.5%	14.6%	9.5%
Precip.	280.0	9.2%	24.2%	13.5%	12.5%	23.4%	17.2%
PM2.5	2.8±2.4	**4.9±1.8**	**1.7±2.5**	**3.6±2**	2.8±2	**2±2.2**	**1.6±2.7**
Memb.	Rocky	13.8%	33.3%	20.6%	15.2%	11.7%	5.4%
Precip.	84.3	19.6%	31.7%	16.1%	16.6%	11.3%	4.7%
PM2.5	2.2±2.5	**3.6±2.1**	**2.8±2.2**	2.1±2.3	1.9±2.4	**1.3±2.4**	**0.8±2.4**
10-day clusters							
Memb.	Denali	27.1%	23.4%	6.8%	20.1%	20.3%	2.3%
Precip.	31.4	26.1%	7.8%	2.0%	36.1%	26.6%	1.4%
PM2.5	0.9±2.6	1.0±2.5	**0.7±2.5**	**1.3±2.8**	1.0±2.8	0.8±2.5	1.2±2.6
Memb.	Rainier	20.6%	16.4%	18.5%	20.9%	18.2%	5.4%
Precip.	280.0	17.5%	13.6%	21.6%	17.4%	24.8%	5.1%
PM2.5	2.8±2.4	**3.7±2.3**	2.9±2	2.6±2.3	3.0±2.4	**2.2±2.4**	2.6±3
Memb.	Rocky	21.7%	3.6%	40.9%	20.1%	10.4%	3.2%
Precip.	84.3	26.9%	4.3%	38.8%	17.7%	9.6%	2.7%
PM2.5	2.2±2.5	**3.2±2.1**	2.2±2.5	**2.6±2.3**	**1.8±2.3**	**1.0±2.5**	**1.0±2.3**

Total average PM2.5 concentration and total annual precipitation (cm) are listed under the site name. Bold PM2.5 concentrations are significantly different than the total mean ($P<0.05$).

different than the site's average concentration at $P<0.05$.

There are a few trends that are readily apparent with the PM2.5 data (see Table 3). Mount Rainier and Rocky Mountain have average concentrations of 2.8 and 2.2 $\mu g\,m^{-3}$, respectively. As could be expected, Denali in Alaska has the lowest average concentration at 0.9 $\mu g\,m^{-3}$.

At all three sites, the shortest clusters, A and B, frequently have concentrations that are significantly higher than average (see Table 3). These clusters are indicative of stagnant air masses generally caused by high pressure. These are conditions that favor poorer dispersion and higher levels of pollution (Moy et al., 1994). Conversely, the longest cluster centers, E and F, have significantly lower than average concentrations. This pattern is seen most clearly with the 1 and 5-day clusters.

For the 10-day clusters, there are fewer cases where the PM2.5 concentrations are significantly different than the average. In addition, the range between the lowest average concentration and the highest concentration is smaller (see Table 3). This trend is the most pronounced at Mount Rainier and

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

Sequoia and Glacier National Parks (not presented). The reason for this is the same as the reason the 10-day clusters failed with precipitation trends at Mount Rainier. The primary sources of PM2.5 are within a few hundred kilometers of the sites. The 10-day clusters have a high variance about the cluster mean for the first few hundred kilometers, causing overlap between the member trajectories, thus reducing their ability to represent distinct transport pathways.

Sources of PM2.5 include sea salt, soil dust, wildfires, and volcanoes. Anthropogenic sources include fossil fuel combustion and other industrial sources. Human-induced sources such as agriculture and prescribed burns are also substantial contributors to atmospheric particulate matter (Hidy and Blanchard, 2005). Rocky Mountain and Mount Rainier are located close to major anthropogenic sources.

Rocky Mountain National Park sits on the front range of the Rocky Mountains. To the east lies the Denver-Fort Collins-Colorado Springs urban corridor, as well as agricultural areas (Mast et al., 2003). In the summer, Rocky Mountain is susceptible to upslope winds caused by the temperature contrasts that form between the mountains and lower surroundings. However, as mentioned, the HYSPLIT trajectories are not likely to represent sub-grid processes. Although cluster centers A and B for the 1-day trajectories are aligned north and south (see Fig. 3), it must be remembered that the spread of trajectories as depicted by the variance about the cluster centers is a better representation of the total path of the cluster. The dominance of summertime flow within these clusters and their significantly higher than average PM2.5 concentrations suggests that these clusters represent PM2.5 from the more polluted eastern side of the Rockies (see Tables 2.2 and 3). The 5 and 10-day clusters follow a similar trend. In both cases, cluster A represents trajectories with flow from the east and has the highest average concentration. Clusters E and F are indicative of marine flow and lower than average concentrations (see Figs. 2.2 and 2.3).

At Rocky Mountain, we have seen the effects that synoptic patterns and emissions have on PM2.5 concentrations. Though both of these factors play major roles in atmospheric pollution, they are often difficult to separate. For example, PM2.5 concentrations typically peak in the summer (Malm et al., 1994), while at the same time clusters with the highest concentrations are composed mostly of summer trajectories. It is possible that these clusters have higher than average concentrations because they contain a disproportionate number of summer days. To test this, we calculated the seasonal average PM2.5 concentration at each of the sites and the seasonal average concentrations within each of the individual clusters. We used a *t*-test to determine whether the seasonal average within each cluster is significantly different than the total seasonal average. Lack of a significant difference would imply that a cluster's concentration reflects only the seasonality of that cluster.

We evaluated the seasonal effects at all sites, and present the results for the 1-day clusters at Mount Rainier (see Fig. 1.1), which have the most pronounced trend. Fig. 5 depicts the concentrations, where the black line represents the average concentration for each season, and the colored lines represent the average seasonal concentrations for individual clusters. A large dot indicates if the seasonal concentration for the given cluster is significantly different than average. Clusters A and B have the highest concentrations, averaging 4.9 and 4.5 $\mu g\,m^{-3}$, respectively (see Table 3). These clusters are short, and encompass the air above the urban Puget Sound area. For all seasons, cluster A is significantly higher than average. With the exception of summer, cluster B is also significantly higher than average. The inverse is true with clusters E and F (see Fig. 5). All seasonal averages for these two clusters are significantly lower than average. This is consistent with rapid transport of marine air that has little influence from local sources. Clusters

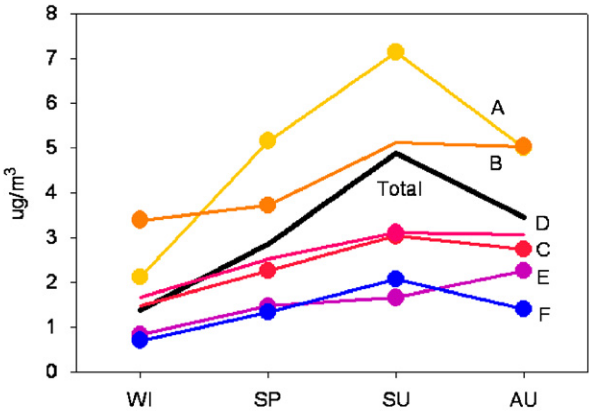

Fig. 5. Breakdown of PM2.5 concentrations by season and by cluster for Mount Rainier. Colors are coordinated to match the cluster color scheme of Fig. 1. The black line is the cumulative seasonal distribution. Dots indicate statistically significant difference.

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

C and D represent a mix of trajectories. Cluster C is significantly lower than average for all seasons but winter, and cluster D is only lower than average in the summer. The PM2.5 concentrations of these two clusters lie between that of the marine and local air, suggesting combined influence. Jaffe et al. (2005) found similar results for PM2.5 at Crater Lake, OR, where trajectories were classified by direction and length. They found local trajectories had the highest concentrations for all months of the year, marine trajectories the lowest, and mixed continental/marine had intermediate values (Jaffe et al., 2005).

Clearly, cluster analysis is capable of segregating pollution data into separate and distinct groups. One aspect of PM2.5 concentration that clusters do not take into account are short-term, high concentrations events; either local as in forest fires, or distant as in Asian/trans-Pacific plumes (Husar et al., 2001; Jaffe et al., 1999, 2003). These events may contribute to high concentrations, but with an aggregate method like clustering, individual events are lost in the big picture of averaged trajectories and concentrations. Even Asian dust, which is thought to be transported to the Western US from Spring though Autumn (VanCuren and Cahill, 2002), cannot be discerned by this method. The cluster centers made with the 1 and 5-day trajectories do not reach Asia, while many of the 10-day trajectories that reach Asia on average contain lower PM2.5 concentrations than air masses that recently spent time in the continental boundary layer. In nearly all cases, the longest cluster F had the lowest average PM2.5 concentration, significantly less than the regionally polluted cases.

4. Conclusions

We calculated and clustered 7 years worth of daily trajectories. Clusters calculated with shorter lengths are better than the 10-day clusters at differentiating precipitation amounts. At each site, shorter clusters tend to occur in the summer, and longer clusters in the winter.

The highest PM2.5 concentrations are associated with the slow moving, regionally polluted clusters and the lowest concentrations belong to the fast moving, typically marine clusters. As with the precipitation trends, the 1 and 5-day clusters had concentrations that were significantly different than average more frequently than the 10-day clusters. At Mount Rainier, a strong seasonal component existed within the clusters. The seasonal average PM2.5 concentrations within each cluster differed from the seasonal average of the entire site.

Overall, clusters generated with 1-day trajectories performed best at differentiating the clusters by precipitation and concentration. Next best were the 5-day trajectories. The 10-day trajectory clusters predominately consisted of flow representative of planetary scale motion, making them less representative of regional transport.

Acknowledgments

This project was funded by the National Park Service as part of the Western Airborne Contaminants Assessment Project (WACAP). We would like to thank Chris Owen for the clustering algorithm, and accompanying help.

References

Abdalmogith, S.S., Harrison, R.M., 2005. The use of trajectory cluster analysis to examine the long-range transport of secondary inorganic aerosol in the UK. Atmospheric Environment 39, 6686–6695.

Baron, J., Denning, A.S., 1993. The influence of mountain meteorology on precipitation chemistry at low and high elevations of the Colorado Front Range, USA. Atmospheric Environment 27A (15), 2338–2349.

Brankov, E., Henry, R.F., Civerolo, K.L., Hao, W., Rao, S.T., Misra, P.K., Bloxam, R., Reid, N., 2003. Assessing the effects of transboundary ozone pollution between Ontario, Canada and New York, USA. Environmental Pollution 123, 403–411.

Buchanan, C.M., Beverland, I.J., Heal, M.R., 2002. The influence of weather-type and long-range transport on airborne particle concentrations in Edinburgh, UK. Atmospheric Environment 36 (34), 5343–5354.

Daly, G.L., Wania, F., 2005. Organic contaminants in mountains. Environmental Science and Technology 39 (2), 385–397.

Doesken, N.J., Pielke, R.A., Bliss, O.A.P., 2003. Climate of Colorado. Climatography of the US. Colorado Climate Center, Atmospheric Science Department, Colorado State University, Fort Collins, CO.

Dorling, S.R., Davies, T.D., 1995. Extending cluster analysis—synoptic meteorology links to characterize chemical climates at six northwest European monitoring stations. Atmospheric Environment 29 (2), 145–167.

Dorling, S.R., Davies, T.D., Pierce, C.E., 1992. Cluster analysis: a technique for estimating the synoptic meteorological controls on air and precipitation chemistry—method and applications. Atmospheric Environment 26A (14), 2575–2581.

Draxler, R.R., Hess, G.D., 2004. Description of the HYSPLIT_4 Modeling System. NOAA Air Resources Laboratory, Silver Spring, MD.

Hageman, K.J., Simonich, S.L., Campbell, D.H., Wilson, G.R., Landers, D.H., 2006. Atmospheric deposition of current-use and historic-use pesticides in snow at national parks in the

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049

Western US. Environmental Science and Technology 40 (10), 3174–3180.

Harris, J.M., Kahl, J.D.W., 1994. Analysis of 10-day isentropic flow patterns for Barrow, Alaska: 1985–1992. Journal of Geophysical Research 99 (D12), 25845–25855.

Harris, J.M., Draxler, R.R., Oltmans, S.J., 2005. Trajectory model sensitivity to differences in input data and vertical transport method. Journal of Geophysical Research 110, D14109.

Hidy, G.M., Blanchard, C.L., 2005. The midlatitude North American background aerosol and global aerosol variation. Journal of the Air & Waste Management Association 55 (11), 1585–1599.

Husar, R.B., Tratt, D.M., Schichtel, B.A., Falke, S.R., Li, F., Jaffe, D., Gasso, S., Gill, T., Laulainen, N.S., Lu, F., Reheis, M.C., Chun, Y., Westphal, D., Holben, B.N., Gueymard, C., McKendry, I., Kuring, N., Feldman, G.C., McClain, C., Frouin, R.J., Merrill, J., DuBois, D., Vignola, F., Murayama, T., Nickovic, S., Wilson, W.E., Sassen, K., Sugimoto, N., Malm, W.C., 2001. Asian dust events of April 1998. Journal of Geophysical Research—Atmospheres 106 (D16), 18317–18330.

Jackson, P.L., 1993. Atlas of the Pacific Northwest: Climate 48-57. OSU Press, Corvalis, OR.

Jaffe, D., Anderson, T., Covert, D., Kotchenruther, R., Trost, B., Danielson, J., Simpson, W., Berntsen, T., Karlsdottir, S., Blake, D., Harris, J., Carmichael, G., Uno, I., 1999. Transport of Asian Air Pollution to North America. Geophysical Research Letters 26 (6), 711–714.

Jaffe, D., McKendry, I., Anderson, T., Price, H., 2003. Six 'new' episodes of trans-Pacific transport of air pollutants. Atmospheric Environment 37, 391–404.

Jaffe, D., Tamura, S., Harris, J., 2005. Seasonal cycle and composition of background fine particles along the west coast of the US. Atmospheric Environment 39 (2), 297–306.

Kahl, J.D.W., Martinez, D.A., Kuhns, H., Davidson, C.I., Jaffrexo, J.L., Harris, J.M., 1997. Air mass trajectories to Summit, Greenland: a 44 year climatology and some episodic events. Journal of Geophysical Research 102 (C12), 26861–26875.

Keim, B.D., Meeker, L.D., Slater, J.F., 2005. Manual synoptic climate classification for the East Coast of New England (USA) with an application to PM2.5 concentration. Climate Research 28, 143–154.

Landers, D.H., Erway, M.M., Kent, M., Blett, T., Simonich, S., Geiser, L., Schreck, C., Taylor, H.E., Campbell, D.H., Jaffe, D., 2003. Western Airborne Contaminants Assessment Project Research Plan. EPA/600/R-03/035. US Environmental Protection Agency, Office of Research and Development, NHEERL, Western Ecology Division. Corvallis, OR.

Mahura, A.G., Jaffe, D.A., Andres, R.J., Merrill, J.T., 1999. Atmospheric transport pathways from the Bilibino nuclear power plant to Alaska. Atmospheric Environment 33, 5115–5122.

Malm, W.C., Sisler, J.F., Huffman, D., Eldred, R.A., Cahill, T.A., 1994. Spatial and seasonal trends in particle concentration and optical extinction in the US. Journal of Geophysical Research 99 (D1), 1347–1370.

Mast, M.A., Campbell, D.H., Ingersoll, G.P., Foreman, W.T., Krabbenhoft, D.P., 2003. Atmospheric Depositions of Nutrients, Pesticides, and Mercury in Rocky Mountain National Park, Colorado, 2002. Reston, VA, US Geological Survey, 15p.

Moy, L.A., Dickerson, R.R., Ryan, W.F., 1994. Relationship between back trajectories and tropospheric trace gas concentrations in rural Virginia. Atmospheric Environment 28 (17), 2789–2800.

Owen, R.C., 2003. A Climatological Study of Transport to the PICO-NARE Site Using Atmospheric Backward Trajectories. Civil and Environmental Engineering. Michigan Technological University, Master of Science, Houghton, MI.

Sirois, A., Bottenheim, J.W., 1995. Use of backward trajectories to interpret the 5-year record of PAN and O_3 ambient air concentrations at Kejimkujik National Park, Nova Scotia. Journal of Geophysical Research 100 (D2), 2867–2881.

VanCuren, R.A., Cahill, T.A., 2002. Asian aerosols in North America: frequency and concentration of fine dust. Journal of Geophysical Research 107 (D24), 4804.

Wang, Y.Q., Zhang, X.Y., Arimoto, R., Cao, J.J., Shen, Z.X., 2004. The transport pathways and sources of PM10 pollution in Beijing during spring 2001, 2002, and 2003. Geophysical Research Letters 31 (14), L14110.

Please cite this article as: Hafner, W., et al., Analysis of rainfall and fine aerosol data using clustered trajectory analysis for National Park sites in the Western US. Atmospheric Environment (2007), doi:10.1016/j.atmosenv.2006.11.049