Reducing sampling error in Rocky Mountain atmospheric carbon dioxide time series to improve flux retrievals

Bjørn-Gustaf J. Brooks¹*, Ankur R. Desai¹, Britton B. Stephens²

1. University of Wisconsin-Madison, Madison, Wisconsin

2. National Center for Atmospheric Research, Boulder, Colorado

1. Introduction

Establishing accurate CO₂ fluxes by atmospheric inverse approaches in mountainous terrain depends critically on the coverage of observed CO_2 concentrations. The Regional Atmospheric Continuous CO₂ Network in the Rocky Mountains (Rocky RACCOON, www.raccoon.ucar.edu) provides open access to atmospheric carbon dioxide measurements covering the central and southern U.S. Mountain West, a region where the exchange of carbon dioxide between the atmosphere and biosphere is not well represented by models. In terrain, measurements of CO₂ are often biased by incomplete mixing and air masses that are locally representative. Here we evaluate two simple and two sophisticated methods of filtering Rocky Mountain CO₂ concentration data in terms of their ability to reduce sampling error by screening locally biased and transient measurements while retaining data sampled during intervals indicative of regionally well-mixed air.

RACCOON is a continuously running array of mountaintop CO_2 towers that bridges an important geographic gap in coverage of Mountain West atmospheric carbon cycling (see Figure 1). Data sets from the still-growing RACCOON network range from August 2005 through present and were in place prior to disturbance events including a widespread mountain pine beetle infestation of the lodgepole pines at Fraser Experimental Forest and the surrounding region. The RACCOON data set captures these events, however it is not yet possible to distin-

guish their impacts in the Mountain West either directly through CO_2 measurements or inverse methods. Our objectives through this work are to determine if spatial coherence and simple evaluation of CO_2 time series can be used to identify local or regional air masses. Also, how do simple subsetting/filtering methods compare to statistical CO_2 filters with respect to variances, growth rates, and 'flagged data'? And finally, what do these filters suggest about the use of mountaintop CO_2 observations to estimate regional fluxes through inverse models?

2. Background

Atmospheric budgets and inversions are preferred in mountainous regions because measuring CO_2 flux using contemporary eddy covariance techniques in mountainous terrain can lead to regionally unrepresentative results. This is particularly true for a significant fraction of the United States land surface where the requisite eddy covariance constraints (e.g. fully turbulent air mass, uniformly horizontal landscape) are confounded by the thermally and terrain induced flows associated with complex terrain.

In theory inverse methods, such as the Carbon-Tracker tracer-transport system, are better able to retrieve CO_2 fluxes because they assimilate observed CO_2 concentrations across a network of *in situ* sensors and use those data to update flux estimates. These data assimilation schemes are potentially very useful accounting schemes for estimating the regional variability of atmospheric CO_2 (e.g. Peters and others 2010), and in addition they provide a means to carry out simulated experiments (*i.e.* Observing System Simulation Ex-

^{*}Corresponding author address: Bjorn Brooks, Univ. of Wisconsin-Madison, 1225 W. Dayton St., Center for Climate Research, Madison, WI 53706-1612; e-mail: bjorn@climatemodeling.org

Figure 1: Map of RACCOON domain and contrasting seasonal CO_2 cycles. A. The complimentary positioning of the mountaintop network of autonomous CO_2 concentration towers from the RACCOON is shown with reference to NOAA'S GLOBALVIEW continuous CO_2 tower network. B. gives a comparison between a baseline seasonal CO_2 curve from MLO shown by the black area filled with cross hachures, and three select towers from the RACCOON domain, Fraser (FEF), Niwot (NWR), and Storm Peak (SPL). All curves represent the average monthly means from years 2006 through 2009.



periment's) for determining ecosystem response to carbon cycle perturbations as well as the sensitivity of an observation network to such change.

The issues however, with inverse methods over complex terrain are the denseness of the observations, their coverage, and their representativeness. An illustration of the potential pitfalls of measuring CO_2 in the mountains appears in Figure 2. It shows

Table 1: Location and characteristics of the RAC-COON towers.

Tower	lat., l	on.	Setting
EFS	38.80 N, 1	09.21 W	marginal plateau
FEF	39.91 N, 105.88 W		alpine valley
RBA	36.46 N, 109.10 W		mountaintop
SPL	40.45 N, 106.73 W		mountaintop
HDP	40.56 N, 1	11.65 W	mountaintop
NWR	40.05 N, 1	05.58 W	mountaintop
	Elevation	tower	Year
	(msl)	ht. (m)	Installed
EFS	1280	39	2007
FEF	2745	18	2005
RBA	2982	22	2007
SPL	3210	9	2005
НПР		40	0000
	3351	18	2006
NWR	3351 3523	18 5	2006 2005

 CO_2 measurements made during multiple downward spirals over the RACCOON domain during the Airborne Carbon in the Mountains Experiment (ACME 2007, Ahue et al. 2009). Although mountaintop towers are often able to sample well-mixed air that is fairly uniform throughout the boundary layer, they also sample poorly mixed air. The challenge is to discern between well-mixed and local air without expensive daily airborne CO_2 profiles or very tall towers.

There are six towers within the RACCOON network that span Colorado, Arizona and Utah (Table 1). The Fraser Experimental Forest tower (FEF) is located about 100 km from Denver and is situated within a high elevation subalpine coniferous forest study area. FEF is quite different from all other RACCOON towers because it sits in a shallow topographic basin about 1 km across, and is surrounded by higher terrain of approximately 50 m relief. FEF lies near the lowest part of the basin where pools of poorly mixed air can persist before being flushed by valley winds. The FEF tower was in place prior to the onset of the mountain pine beetle infestation of the lodgepole pines throughout the region.

Figure 2: Example sampling error. Comparison of composite plots that represent multiple downward spirals from the 2007 ACME airborne campaign (Ahue et al. 2009) over the RACCOON domain reveals a strong contrast between morning (upper pane) vertical CO_2 profiles and afternoon (lower pane). In the morning profile a large toe of CO_2 appears at the bottom of the plot illustrating the potential for sampling error due to pooling of CO_2 from nocturnal respiration, which is flushed-out by the afternoon.



The Niwot Ridge T-Van site (NWR) sits near the tree-line on Niwot Ridge 5 km and 470 m up the ridge from the Niwot Ridge Ameriflux site. Storm Peak Lab (SPL), Hidden Peak (HDP) and Roof Butte (RBA) are all mountaintop facilities further removed from vegetation that can experience prolonged episodes

above the cloud base especially during winter. Entrada Field Station (EFS) sits much lower in elevation on the edge of the Colorado plateau in a region dominated by canyons, mesas and steep cliffs.

3. Discrete-time CO₂ Filters

Procedures for filtering mountaintop carbon dioxide time series have been in use for several decades (Pales and Keeling 1965, Gillette and Steele 1983, Thoning et al. 1989) shortly after it was recognized that mountaintop CO₂ measurements of well-mixed air can be masked by a variety of phenomena including nocturnal pulses representing local air enriched in CO_2 and by daytime troughs in CO₂ caused by local vegetation and boundary layer covariance (Pales and Keeling 1965, Keeling et al. 1976)- both of which are essential problems when assimilating CO₂ concentration data. We experimented with three different kinds of discrete-time filters and one spatial coherence technique in order to determine the feasibility of extracting synoptic signals with low sampling error from our CO₂ time series. The first method (coherence) is a way of subsetting simultaneous observations recorded at multiple towers based on the coherence of their CO₂ values. The second method is a detection error filter (DE) that provides a rapid way of excluding observations based on simple statistics of synchronous observations across tower inlet heights. The final two methods represent more sophisticated means of filtering. The third is a statistical interpolation filter (SI) based on the filter used for NOAA's Mauna Loa tower, MLO (Thoning et al. 1989). Our SI filter is nearly identical except that we do not use the additional constraint of a low-pass filter. The last filter is a derivative of Tukey's weighted median smoother (Tukey 1974), except that our method determines a cutoff range though a backward looking window that varies according to season and the secular rise in atmospheric CO_2 concentration.

3.1 Subsetting for Coherent Observations

We used an additional subsetting method, coherence, which is not explicitly a filter but may be considered a null test. We also use coherence as a method of quantifying network heterogeneity in CO_2 observations. Here, coherence represents the likelihood that two towers sample similar air masses, or as the similarity in tower footprints. We occasionally include coherent subsets when comparing the statistics of filtered subsets.

Coherence is the proportion of simultaneous observations between two towers that are similar to within 2 PPM. The coherent proportion is calculated as the fraction of the total number of simultaneous observation points in x(t). If all observations meet the conditions of coherence then the coherence ratio would be 1. There is an implicit assumption that the towers are proximal enough for their footprints to overlap, and therefore we only consider the tower coherence for three towers (FEF, NWR, SPL).

3.2 Method One: Detection Error Filtering Across Inlet Heights

Our detection error (**DE**) filter is a combined sample error and detection error filter that is a fast routine for identifying discordant measurements across tower inlet heights. The DE filter operates by rejecting CO_2 observations that have both excessive hourly standard deviations and observations that occur when a substantial CO_2 gradient exists across the top two inlet heights. Such bias is caused for example by local flows or poorly mixed air. A process similar to the detection error method was used as a first pass filter for mountaintop CO_2 observations at Mauna Loa in Hawaii by Keeling et al. (1976) and later by Gillette and Steele (1983) to reject 'contaminated' mountaintop flask measurements at Niwot Ridge.

The DE filter relies strictly on *a priori* knowledge of acceptable variance conditions. Given the sampling precision of the AIRCOA instrumentation used here we set a 1 PPM limit on the hourly standard deviation along with a 0.5 PPM limit on the difference in CO₂ concentration between the top two inlet heights. This can be described in terms of time series signals. The related discrete-time signal x(t) is extracted from the original signal X(t) where the hourly mean standard deviation at the top inlet height σ_{x_h} is less than 1 PPM and the absolute difference in CO₂ between the top two inlet heights $|X_h(t) - X_{h-1}(t)|$ is less than 0.5 PPM.

3.3 Method Two: Iterative Filtering of Outliers from a Fitted Polynomial

A more sophisticated statistical interpolation (SI) method for rejecting outliers was applied by Thoning et al. (1989) to produce a subset of mountaintop observations taken from Mauna Loa that represented background (*i.e.* baseline) CO₂. The SI filter was used in combination with a low-pass spectral filter to remove all but the most subtle changes in the ambient diurnal and seasonal harmonic of the CO_2 curve. We developed a variant of this in order to extract a signal from Rocky Mountain CO₂ observations that would be representative of regional differences within the US Mountain West at scales smaller than were necessary for the remote mountaintop observations from Mauna Loa. The principal differences of this SI filter are that: it does not use low-pass frequency filter (for example through Fourier decomposition), and second, our method only rejects the largest outlier with each iteration over the time series as opposed to removing several. The intent is to preserve as much of the regional signal as possible without excessive smoothing of the data.

Our SI filter works by passing a ten day sliding window over the original discrete-time signal X(n)to create subset x(n) that is centered at n and consists initially of samples $X_i(n)$ for $i \le 15$ samples (the 15 hourly CO₂ observations belonging to day n and excluding hours $11, 12, \ldots, 19$). For each window a cubic spline S(X) is fitted through the daily means of $X_i(n)$. The filter then removes the hourly observation X_i at n with the largest residual from spline curve S(X).

The window advances across all *n*'s, re-fitting a new ten day spline with each new window and rejecting no more than one observation per day with each iteration over the entire time series. After no more than 14 iterations (the maximum number of hourly observations that can be rejected for each day), or when the standard deviation of all daily means is less than 0.5 PPM (*e.g.* σ of $X_1(n), X_2(n), \ldots < 0.5$) then the excluded daytime observations X_j from the original time series that are within 0.5 PPM of the final refitted spline curve are incorporated back in to form the final subset.

3.4 Method Three: Filtering of Outliers Using a Weighted Median Smoother

An effective method that has been used in many signal filtering applications is the weighted median smoother (cf. Tukey 1974). Because our intent is not necessarily to smooth but to reject CO_2 observations that are not regionally representative when present in the data, we have modified Tukey's method. We use a Weighted Median Filter (**WM**) that rejects observations when their residuals from the daily median is in excess of the summed and weighted inter-day variance for the previous two weeks.

The WM filter advances a backward-looking window over the original discrete-time sequence X(n)and samples observations to comprise the related subset x(n). The envelope of acceptable CO₂ values at moment n (*i.e.* day n) is centered about the median value $\widetilde{X}(n)$. The limits of the envelope are dependent on the sum of the residuals $[\widetilde{X}(n) - \widetilde{X}(n-1)]$ between each day's median over the previous fifteen days in a weighted geometric series. Thus subset x(n) is composed of those elements X_i from the original time series that do not exceed the limit.

3.5 Descriptive Statistics

Normalized Sum of Squares

To find the total space-time variability within RAC-COON data we compute a normalized total sum of squares as the squared difference between each observation and the grand mean that is scaled by the number of observations. The grand mean \overline{X}_G is the mean of all samples from each tower \overline{X}_i weighted by the number of observations in that sample n_i .

We computed the normalized total sum of squares to a subset of the three most proximal RACCOON towers (FEF, NWR, SPL are all within a 50 km radius) in order to capture the total space-time variation. Thus the total sum of squares SS_N is an expression of the variability of RACCOON CO₂ values that is scaled by the degrees of freedom for each tower. It is used as a basis for comparison of the dispersion between different collections of RACCOON data, for example daytime vs. nighttime or between different filtering methods.

4. Results

We first investigated possible trade-offs between improved representativeness in filtered subsets and loss of information content. Table 2 provides a comparison of the complete set of observations to subsets from our three filter methods in order of the proportion of retained observations. The DE filtered subset retained the largest fraction of original observations (0.777) and the most space-time variability ($SS_N = 17.1 \text{ PPM CO}_2$). The WM subset was more exclusive, retaining about two-fifths of all RACCOON observations (0.423) with variability (SS_N) of 9.2 PPM CO₂. The SI subset was the most exclusive (retained fraction: 0.343) and least variable with SS_N of 8.2 PPM CO₂. Collectively these represent a predictable linear association between the removal of observations and reduction in variability.

Although SI and WM methods filtered similar proportions of observations (42.3%,34.3%), the proportion of agreed observations (i.e. the proportion of total observations that occurred in both WM and SI subsets) only about half or 21.2% of the complete set of observations.

Table 2: Filter statistics for the complete set of observations (CS) and filtered subsets (DE, WM, SI) across years 2005-2010. Normalized sum of squares (SS_N) represents average space-time variability. These show a relationship between decreased retained fraction and decreased variability (SS_T) of the subset.

Set	Retained Fraction	Grand Mean	SS_N
CS	1.000	388.5	50.2
DE	0.777	387.1	17.1
WM	0.423	387.5	9.2
SI	0.343	387.5	8.2

We were also interested in the coherence of observations at different times of day, different seasons, and before and after filtering as a means to determine whether filtering made the tower data more similar. We computed coherence as the proportion of time that proximal towers (FEF, NWR, SPL) had CO₂ values similar to within 1 PPM $(\pm 0.5 \text{ PPM})$. Because coherent CO₂ observations could also be the result of chance, it is important to note that we are using coherence primarily as a measure of similarity comparison between tower observations, not as a method to filter observations. The coherence statistics listed in Table 3 indicate that subsets of 'afternoon' and 'summertime' observations show little difference from the complete set (0.212/0.223:0.201). However, nocturnal observations are 25% more likely (0.264:0.212) to be coherent than afternoon, while wintertime observations are 16% more likely (0.259:0.223) to be coherent than summertime. Past studies (Pales and Keeling 1965, Thoning et al. 1989, de Wekker et al. 2009) have excluded afternoon observations on the basis that local vegetation can significantly bias CO₂ measurements as photosynthesis depletes the local air of CO₂. This metric indicates that nighttime and wintertime are about one-fifth more likely to be coherent, and perhaps sampling well-mixed air.

Table 3: Network similarity by season and time of day. Coherence is similarity in CO_2 values to within ± 0.5 PPM. A completely coherent network of observations would have a ratio of 1. Afternoon and nocturnal represent four-hour means centered over 14.00 and 02.00 local time. Summer and winter represent the mean of all observations from June, July, August and December, January, February. Coherence is generally higher in nocturnal and winter subsets suggesting higher probability of measuring well-mixed air.

	CS	DE	SI	WM
All	0.201	0.513	0.538	0.451
Aftern.	0.212	0.403	0.693	0.315
Noctur.	0.264	0.506	0.411	0.332
Summer	0.223	0.568	0.242	0.252
Winter	0.259	0.409	0.354	0.312

We also examined differences in the mean annual growth rate in CO_2 and its dependent on the collection of observations used. In RACCOON data for years with CO_2 growth rates above 2 PPM/yr all filtered subsets reduce the apparent growth rate by 4-12% (*c.f.* years 2006, 2008 in Table 4). Figure 4

Figure 3: The cost of coherent observations is illustrated for each filter and its subsets (All, Afternoon, Nocturnal, etc.). A perfectly coherent network would be comprised of observations all within ± 0.5 PPM and would lie along the top of the plot. DE subsets for example have both a high coherence and high proportion of retained observations.



Proportion of Retained Observations

illustrates that there is a strong agreement (*i.e.* a small spread in grand mean values) between all filtered subsets.

Table 4: Differences in mean annual growth rates between filtered subsets. Notice that all 2006, 2008 filtered subsets have larger growth rates than the complete set.

Year	All	DE	SI	WM
2006	2.10	2.19	2.42	2.28
2007	1.71	2.10	1.69	1.63
2008	3.10	3.27	3.40	3.48
2009	0.98	0.92	0.76	0.69

4.1 Synoptic Sensitivity Case Studies

In order to diagnose each filter's capacity to enhance the regional representativeness of filtered subsets we examined several cold front system case studies. Frontal systems are often associated with large-scale meteorological changes in temperature Figure 4: Mean annual CO_2 growth rates by year. Growth rates of filtered subsets corresponding to columns DE, SI, WM from Table 4 are plotted against the rates of their complete counterparts corresponding to the 'All' column. Note that growth rates along the vertical axis greater than 2 PPM/year lie below the 1:1 line



and humidity that bear upon the exchange of carbon dioxide between the biosphere and atmosphere. We used multiple frontal passage case studies as a diagnostic test to compare filtering effectiveness.

To standardize our protocol we developed a supervised automated routine that identifies prolonged troughs in barometric pressure that are coupled to decreases in temperature, humidity, and abrupt wind direction shifts. Figure 5 shows the pressure, temperature, and wind conditions associated with a summertime frontal system that passed over NWR in 2007.

There are two important features in the CO_2 data during this frontal system that make this a useful case. The first is a feature relevant to monitoring the regional carbon cycle- the pronounced step-like increase in CO_2 values occurring between -46 and -24 hours prior to the barometric pressure minimum (see lower two panels in Figure 5). Given that a front that influences regional photosynthesis through lower temperatures and reaction rates, the increase in observed CO_2 concentrations probably represents decreased CO_2 uptake in the upwind direction. Such synoptic events impart important inFigure 5: Synoptic case study figure of one case study frontal system and its effect on atmospheric CO_2 . Hours relative to the front appear on the top of each plot, and time of day listed along the bottom.



formation about the regional carbon cycle response and can be used to test the sensitivity of a filter.

We also examined a feature that contributes confounding information to the regional carbon cycle and should be rejected by a filter. Pulses of nocturnal CO_2 are typically localized transient events associated with wind direction shifts can be identified as either sharp increases or decreases in CO_2 concentration that last several hours or less. Figure 5 shows two such pulses that occur at about 72 and 48 hours prior to the frontal passage.

Over a set of 5 selected front systems from 2005 through 2010 the DE filter retained 92% of observations (369/400), WM 36% (145/400), SI 37% (146/400), and coherent subsetting 52% (206/400) making the DE method the least selective filter for these case studies. In addition the DE filter did not reject any of the 7 nocturnal pulses, but preserved nearly all of the daytime troughs in CO₂ (see hours -72, -48 in upper right panel of Figure 5). Coherent subsetting rejected 3 of 7 nocturnal pulses, but retained much of the diurnal CO₂ cycle including strong daytime troughs that are are probably local and not representative of well-mixed regional air.

The SI and WM filters similarly rejected all 7 nocturnal pulses while retaining the synoptic step-like shifts in CO_2 , which resulted in similar mean values for their subsets. Upon closer inspection of the individual values however, the SI and WM filters only share about half of the same observations. That is, the proportion of observations that occurred in both WM and SI subsets was only about half or 21.2% of the complete set of observations.

5. Discussion

Observed carbon dioxide concentrations collected from continuously sampling mountaintop towers must be filtered in order to lower sampling error caused by local events that represent incompletely mixed air masses. This study evaluated two simple and two sophisticated methods for filtering observed CO₂ measurements to ascertain their ability to reject observations that do not contribute useful information about the regional carbon cycle, while preserving those elements that are important to

monitoring synoptic scale variability.

Our first objective was to determine if spatial coherence and simple evaluation of CO_2 time series can be used to identify local or regional air masses. Although the general statistics of coherent and detection error (DE) filtered subsets indicate that they retain one-half to four-fifths of the total observations, along with substantial space-time variability (SS_N) all without biasing the distribution of the data, results from case study synoptic events clearly show that both of these methods do not adequately reject observations clearly indicative of poorly mixed air masses such as nocturnal pulses and exaggerated daytime troughs in CO_2 .

Statistical interpolation (SI) and weighted median (WM) filtering reject roughly one-half to three-fifths of all mountaintop observations from RACCOON and preserve the least total variability. However, these SI and WM subsets far better represent synoptic-scale carbon cycle changes. All nocturnal pulses are rejected and large fluctuations in the diurnal cycles are consistently filtered, while distinct jumps in CO_2 concentration that are linked to synoptic weather/temperature changes are retained in SI and WM filtered subsets.

Filtering RACCOON data has the effect of slightly decreasing the mean annual CO₂ growth rate for years that exceed 2 PPM indicating that unfiltered data may exaggerate growth for such years. SI and WM filtered subsets also have a total variability about one-fifth that of the complete set and one-half that of DE filtered subsets indicating that current inverse methods may be assimilating data with a total carbon cycle variability that is a factor of two too large. Also such methods if they rely on simple filtering methods such as coherence or DE filtering may be including large proportions of original observations, but those observations are likely to include fluctuations in atmospheric CO₂ that are indicative of local events such as nocturnal pulses that co-mingled with legitimate regional CO₂ changes case by synoptic frontal systems.

Acknowledgments

This work was supported by NOAA CPO, grant no. NA09OAR4310065. The authors would also like to thank Kurt Chowanski (Institute of Arctic and Alpine Research, University of Colorado), Gannet Hallar (Director, Storm Peak Laboratory, Desert Research Institute) for their collaboration and providing meteorological data.

References

- Ahue, W. K., A. R. Desai, S. de Wekker, D. J. Moore, T. L. Campos, B. B. Stephens, R. K. Monson, and D. Schimel, 2009: Uncertainty of regional carbon fluxes and boundary layer heights in complex terrain: The airborne carbon in the mountains experiment 2007. Fall Meeting Supplemental, Abstract B53F–08.
- de Wekker, S. F. J., A. Ameen, G. Song, B. B. Stephens, A. G. Hallar, and I. B. McCubbin, 2009: A preliminary investigation of boundary layer effects on daytime atmospheric co₂ concentrations at a mountaintop location in the Rocky Mountains. *Acta Geophysica*, **57**, 904–922, doi: 10.2478/s11600-009-0033-6.
- Gillette, D. A. and A. T. Steele, 1983: Selection of CO₂ concentration data from whole-air sampling at three locations between 1968 and 1974. *Journal of Geophysical Research*, **88**, 1349–1359, doi: 10.1029/JC088iC02p01349.
- Keeling, C. D., R. B. Bacastow, A. E. Bain-Bridge, J. C. A. Ekdahl, P. R. Guenther, L. S. Waterman, and J. F. S. Chin, 1976: Atmospheric carbon dioxide variations at mauna loa observatory, hawaii. *Tellus*, 28, 538–.
- Pales, J. C. and C. D. Keeling, 1965: The concentration of atmospheric carbon dioxide in Hawaii. *Journal of Geophysical Research*, **70**, 6053–6076, doi: 10.1029/JZ070i024p06053.
- Peters, W. et al., 2010: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations. *Global Change Biology*, **16**, 1317–1337.

- Thoning, K. W., P. P. Tans, and W. D. Komhyr, 1989: Atmospheric carbon dioxide at mauna loa observatory. ii - analysis of the noaa gmcc data, 1974-1985. *Journal of Geophysical Research*, **94**, 8549–8565, doi: 10.1029/JD094iD06p08549.
- Tukey, J. W., 1974: Nonlinear (nonsuperposable) methods for smoothing data. *Conference Record: IEEE EASCON*, 673.