# Regional Haze Rule Natural Level Estimates Using the Revised IMPROVE Aerosol Reconstructed Light Extinction Algorithm

**Final Paper # 48**

**S. A. Copeland**, CIRA/USDA Forest Service, 333 East Main Street, Lander, WY 82520;
**M. Pitchford**, Air Resources Laboratory, NOAA, 755 East Flamingo Rd, Las Vegas, NV 89119;
**R. Ames**, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523

## INTRODUCTION

In 1999 the Environmental Protection Agency (EPA) promulgated the Region Haze Rule (RHR)[1] that established a program to mitigate existing visibility impairment in 156 visibility-protected federal class I areas (i.e. certain large national parks and wilderness areas as specified by the Clean Air Act). RHR implementation guidance requires states to determine the uniform rate of haze reduction that would result in reaching natural haze level from current conditions for the five-year mean of the annual 20% most impaired visibility condition for each protected area.[2] They must also show that the five-year mean of the annual 20% least impaired visibility conditions at each protected area do not degrade. For the RHR, haze levels are expressed using units of deciview, which is a log-transformation of light extinction that results in a more perceptually uniform metric of haze than either light extinction or visual range.[3] Light extinction is estimated from particle speciation data generated by the Interagency Monitoring of Protected Visual Environments (IMPROVE)[4] monitoring network that samples on a one day in three schedule at 110 monitoring sites representative of the class I protected areas. Though states may ultimately justify adoption of an alternative rate of progress in their RHR implementation plans, they must first determine the uniform rate of progress for their state's class I areas.

Haze levels for the 20% worst natural haze conditions need to be estimated in order to determine the 60-year uniform rate of progress for each protected area. EPA provided a default approach[5] for estimating worst natural haze levels that involves a three step process. First the mean natural haze levels are estimated using the original IMPROVE algorithm[6] for estimating light extinction from PM speciation data applied to estimates of natural PM species concentrations developed by John Trijonis[7] and a spatially-interpolated 10-year measured relative humidity data set that provides distinct values for each visibility protected area. The second step is to convert the mean natural light extinction levels for each protected area to deciview values using the simple log-transformation that defines the deciview scale.[3] The final step is to add 1.28 times the assumed standard deviation of the natural haze distribution in deciview units to the mean value from the second step to get an estimate of the 90$^{th}$ percentile that was thought to be a reasonable approximation of the mean of the 20% highest haze levels. Estimates of the

20% least impaired natural conditions were made in a similar way by subtracting 1.28 times the assumed standard deviation of the natural haze distribution in deciview units. The assumed standard deviation of 3dv for eastern class I areas and 2dv for western class I areas was based on the distribution of data at the most pristine locations in the east and west.[8] Following this procedure, EPA provided estimated default mean, worst and best natural haze levels for all of the protected areas as part of their guidance which states could use unless they chose to refine the approach.

The default approach for estimating natural conditions was the subject of technical reviews as a result of its role in determining the uniform rate of progress for the RHR.[9, 10, 11] The Trijonis estimates of typical natural species concentrations which provided one set of values for the eastern U.S. and another set for the western U.S. were based on the limited available PM composition monitoring data in the late 1980s, so were seen as not likely to be representative of all of the visibility protected areas. The statistical approach that calculated the 90$^{th}$ and 10$^{th}$ percentile natural levels for use as estimates of the mean of the 20% best and 20% worst haze levels was shown to be flawed in a number of ways. The IMPROVE algorithm was shown to overestimate light extinction at low levels and to underestimate it at high levels. It also didn't work well for data from coastal monitoring sites because it didn't account for light scattering by sea salt.

The IMPROVE algorithm was revised in 2005 to reduce biases at the high and low light extinction extremes, to incorporate more recent information from the literature, to include extinction from sea salt (an important component at some coastal sites), and to include elevation-specific estimates of Rayleigh light scattering.[12] Many states indicated their intent to adopt the new algorithm for RHR assessment and planning purposes. For consistency in the determination of a uniform rate of progress, the use of the revised algorithm for estimating current haze levels implies the need to use it to estimate natural conditions.

In 2006 the Regional Planning Organizations (i.e. five regional organizations of state, tribal, federal and stakeholder organizations established to promote regionally consistent RHR implementation) established an ad hock committee to establish a uniform approach for the application of the revised IMPROVE algorithm to estimate natural levels and to remedy other flaws in the default methodology where possible. A presentation of the results of their work is available on the IMPROVE web site.[13] In the sections that follow is a description of the revised natural haze estimation methodology, the results of its application, and a discussion of the differences between the default and revised natural haze estimates and uniform rate of progress results.

## METHODOLOGY

There are two fundamental differences between the default and the revised approach for estimating the 20% highest and 20% lowest natural haze levels for each of the visibility protected areas. The first difference involves the use of the revised IMPROVE algorithm for estimating light extinction from PM speciation data in place of the original IMPROVE algorithm. The second difference addresses the flaw in the default approach's use of a statistical method to estimate the highest and lowest natural haze levels. The revised approach adjusts each sample period's species concentration to generate a simulated natural haze distribution with the annual mean for each species being equivalent to the Trijonis estimated natural concentration for that species. The step by step details of this approach are described below.

1. Start with the "Daily Values Including Patched Values" data set from the on the Visibility Information Exchange Web Sites (VIEWS) data summary website[14].
2. For sites with fewer than three valid sample years from the 2000-2004 baseline period, use the substituted data[15] set in place of the standard VIEWS data from step 1.
3. Select records with sample dates from 2000-2004 for each site of interest.
4. Retain only records from "valid" years based on EPA RHR guidance[2].
5. Retain only records that have "valid" sample codes for each of the seven aerosol species (i.e. sulfate, nitrate, organic carbon, elemental carbon, fine soil, coarse mass, and sea salt) needed to reconstruct aerosol extinction using the revised IMPROVE algorithm[12].
6. Discard records which have any patched values from the EPA RHR patching technique[2].
7. Calculate arithmetic means for each of the seven species for each site and each calendar year.
8. Divide the resultant means by the Trijonis-based EPA default natural conditions estimates[5] for the six species (i.e., all except sea salt, which is included in the revised IMPROVE algorithm but not included in the default estimates). "East" and "West" are defined as east or west of the 98$^{th}$ meridian respectively for purposes of determining which default natural conditions to use. The ratio of baseline annual mean to default natural annual mean becomes the scaling factor for each of the species.
9. In cases where the baseline annual mean concentration of a species is less than the EPA default estimated concentration, the baseline values are retained (i.e. the scaling factor is 1). Similarly, sea salt concentrations have a scaling factor of 1, which defines the baseline condition as the natural condition.
10. Every observation is divided by the appropriate species specific scaling factor, creating a new distribution for each species which has the same annual mean as the EPA default natural annual mean (or the baseline value when a scaling factor of 1 is used) while retaining the shape of the baseline distribution. Figure 1 illustrates the adjustment of species concentration for one site-year.

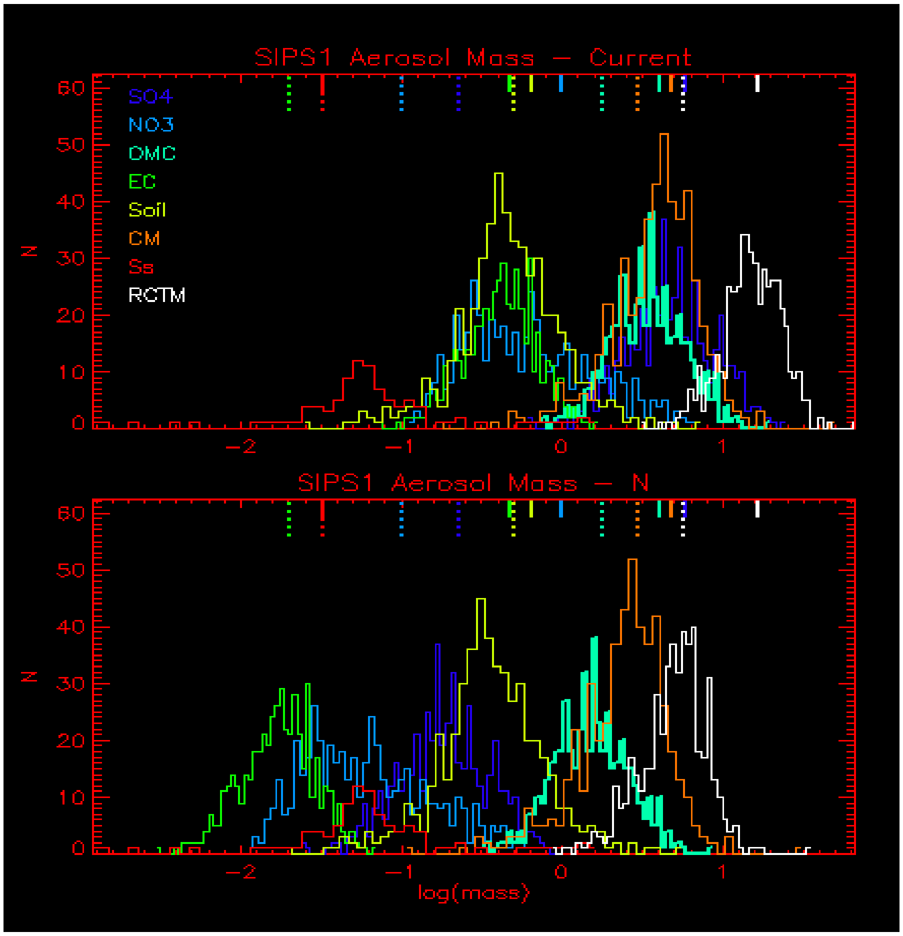


**Figure 1. Frequency distribution of the current (i.e. measured) species concentration for Sipsey Alabama (top) and the species concentrations adjusted to Trijonis-based natural levels (bottom). The use of a logarithmic concentration scale means that the distributions are the same shape for each current and natural species, though translated horizontally. No log transformation is performed in the approach however. The hanging bars at the top of each plot show the mean for each species (see color key) with short solid lines being current and longer dotted lines being adjusted means.**

11. Keeping the observations grouped by the sample date, convert the scaled masses of each species to extinction values using the revised IMPROVE algorithm.
12. Sum the resultant species specific extinctions for each sample date, and add the site specific Rayleigh scattering value.
13. Convert this daily extinction value to a deciview value. For pristine conditions at high elevation sites (i.e. >2200m) these deciview values are sometimes negative. While counterintuitive, this is mathematically appropriate and negative or zero values are retained. The result is up to five years worth of daily deciview values adjusted to be simulated natural haze levels at each IMPROVE site. Figure 2 illustrates the current and natural haze distributions that result from this process.

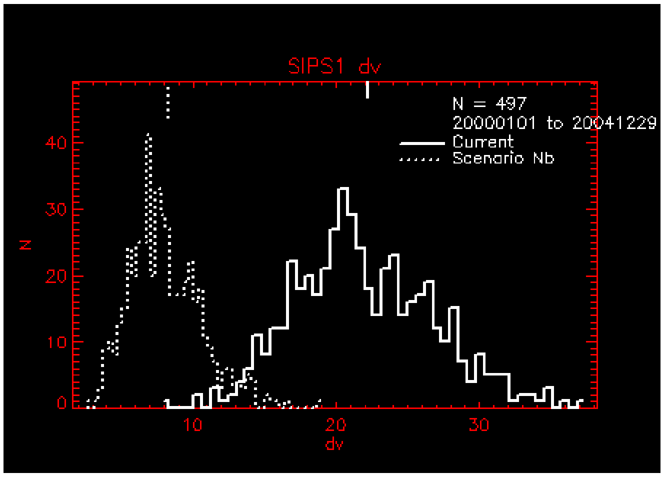


**Figure 2. Frequency distribution of current and adjusted natural haze levels in deciview units for Sipsey Alabama. Notice that the distribution shape is not the same.**

14. Determine the highest and lowest 20% simulated natural deciview days for each year and site in the same way that the highest and lowest 20% deciview days are determined for baseline haze levels. Note that these sample dates will probably not be the same sample dates that were the highest and lowest 20% dates for the baseline calculations.
15. Calculate the arithmetic mean for each year and each species for sample dates which comprise the lowest 20% and highest 20% haze days.
16. Calculate the arithmetic mean across the five annual means for each species including deciview.

The results of these steps are lowest 20% and highest 20% natural conditions estimates for each site with species compositions, and the natural deciview values needed for glide slope calculation. Note that since the deciview transformation is non-linear, the mean deciview value will not equal the deciview value calculated from the sum of the mean aerosol extinction values plus Rayleigh.

## RESULTS

Estimates of the mean, highest 20% and lowest 20% haze levels for each of the IMPROVE monitoring sites representing visibility protected areas is available as the Natural Haze Levels II (version 2) spreadsheet on the VIEWS website[16]. Figure 3 contains contour maps that show the spatial distribution of the EPA default and revised approach estimates for the highest 20% values. Both have the same striking east – west gradient that results directly from the different Trijonis estimates of natural species concentrations for the eastern and western states. The greatest difference between the patterns on the two maps is along the California coast where sea salt contributes significantly to the $PM_{2.5}$ concentration and hence to the haze at coastal monitoring sites. The default approach doesn't account for sea salt while the revised approach explicitly includes it. The natural haze levels from the revised approach at some high elevation sites are lower due to the use of a lower, elevation-dependent Rayleigh light extinction value compared to the default approach. However many other high elevation sites have higher natural haze level estimates using the revised approach (principally those in the Pacific Northwest), probably due to the broader distribution of current and thus simulated natural haze levels associated with a higher incidence of wildfire influence.

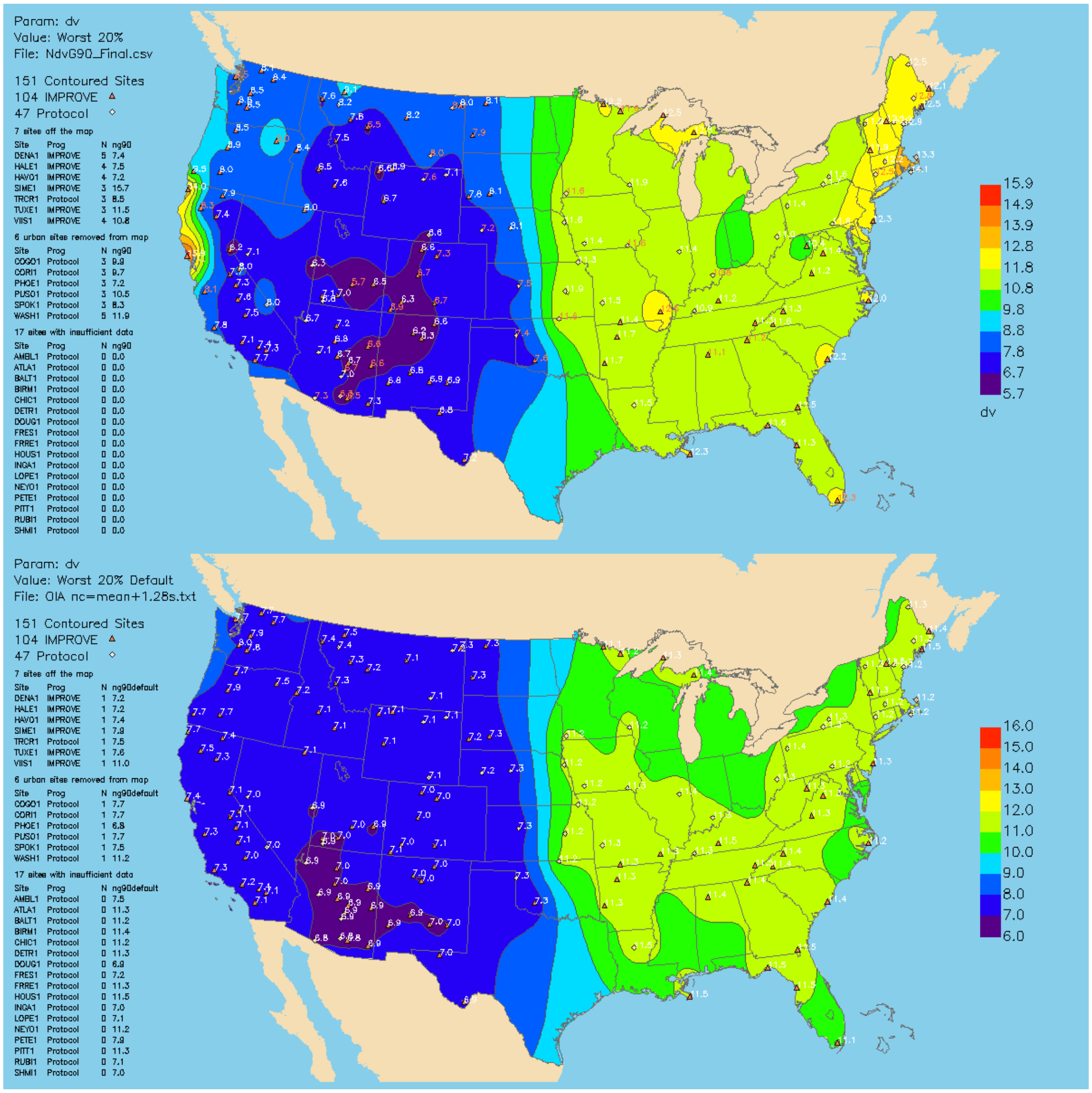

**Figure 3. Contour maps of the revised (top) and default (bottom) natural haze level estimate.**

Perhaps of greater interest is a comparison of the uniform rate of progress values for the revised approach compared to the default approach. Figure 4 shows contour maps of the uniform rate of progress in units of deciview per decade. As with the natural haze level contour maps, the two algorithms yield similar looking spatial patterns, with the greatest rates centered on the Ohio River basin in the East and Southern California in the West. Notice that the revised approach uniform rate of progress values for California coastal sites are not obviously affected by sea salt since its contribution to haze is counted in both the natural and baseline condition estimates.

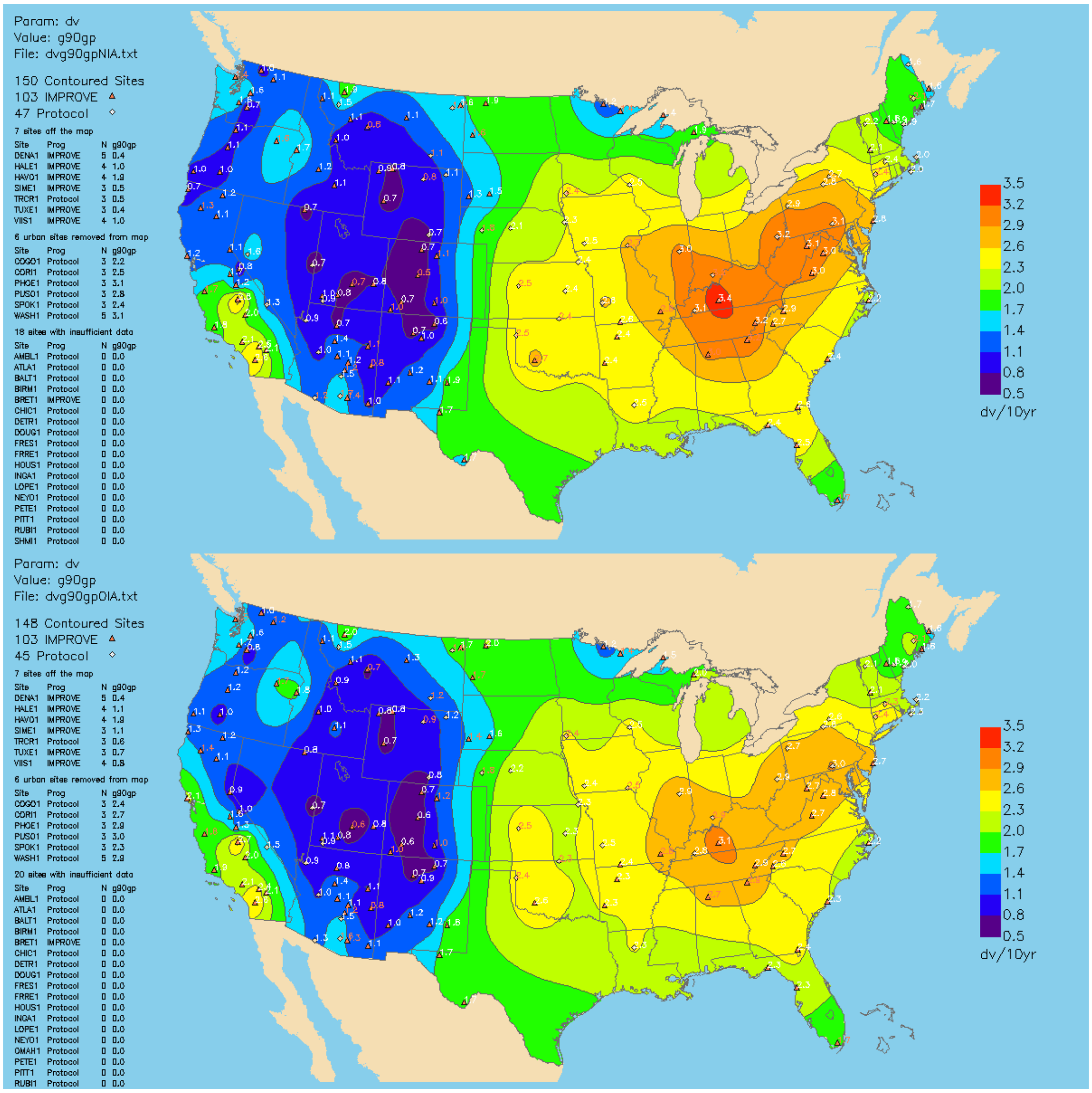


**Figure 4. Contour maps of the 10 year haze reduction in deciview per decade units for a uniform rate of progress for the revised (top) and default (bottom) approaches of estimating natural levels and current conditions.**

## DISCUSSION & CONCLUSIONS

The decision by a number of states to adopt the revised IMPROVE algorithm to estimate haze levels from IMPROVE speciation data resulted in the need for a consistent approach for estimating natural haze levels. In addition to using the new IMPROVE algorithm, the methodology for estimating revised natural haze levels changed the method of estimating the 20% haziest and 20% clearest conditions to avoid the flawed statistical approach that was used to generate the default natural haze levels.

Both the default and revised approaches for estimating natural haze levels rely extensively on the Trijonis East and West estimates of natural species concentration levels, which is undoubtedly why the spatial distributions in Figure 3 show the similar strong gradients along the arbitrarily selected demarcation line between east and west. The only justification for using only two geographically distinct sets of natural PM speciation concentration estimates for all of the visibility protected areas is the lack of estimates for other smaller regions. Tombach[9] shows that up to 15 geographically distinct regions are justified based on similar characteristic of current PM species concentrations, but his work doesn't offer estimates of natural levels for these regions. Global scale modeling may ultimately provide better estimates of natural haze levels on a more spatially and temporally resolved basis, as well as address the related issue of estimating haze from non-U.S. man-made emissions for each visibility protected area.[15]

The inclusion in the revised approach of sea salt with the assumption that all sea salt is from natural sources demonstrates how sample-period-specific monitoring data can be used to refine the estimate of natural levels. However, most PM species are from a combination of natural and man-made sources so some type of attribution analysis would be needed to apportion how much of each species is from either source category on a sample-specific basis. It's unclear whether application of either receptor or air quality simulation modeling would sufficiently improve the accuracy of natural level concentrations over those provided by Trijonis or similar estimates that could be developed for a greater number of regions.

The 60-year schedule of the RHR with its periodic planning and technical review provides the time needed to further expand our knowledge and technology to improve our estimates of natural levels. Ultimately as man-made emissions that contribute to haze in visibility protected areas are reduced, current conditions will more closely match natural haze levels, perhaps making the task of specifying them somewhat easier.

## References

1. 40CFR Part 51: Regional Haze Regulation: Final Rule; *Fed. Regist.* **1999**, 64 (126), 35714-35774
2. *Guidance for Tracking Progress Under the Regional Haze Rule*, EPA-454/B-03-004, September 2003, available at the IMPROVE web site, http://vista.cira.colostate.edu/improve/Publications/GuidanceDocs/guidancedocs.htm (accessed 3/19/2008).
3. Pitchford, M.L.; Malm, W.C., Development and Application of a Standard Visual Index, Atmos. Environ., 28: 5, pp1049-1054, 1994.
4. Interagency Monitoring of Protected Visual Environments (IMPROVE) web site, http://vista.cira.colostate.edu/improve/Default.htm (accessed 3/19/2008).
5. *Guidance for Estimating Natural Visibility Conditions Under the Regional Haze Program*, EPA-454/B-03-005 September 2003, available at the IMPROVE web site, http://vista.cira.colostate.edu/improve/Publications/GuidanceDocs/guidancedocs.htm (accessed 3/19/2008).
6. Sisler, J.F., *Spatial and Seasonal Patterns and Long Term Variability of the Composition of the Haze in the United States: an Analysis of Data from the IMPROVE network:* 1996 CIRA Report to the National Park Service, ISSN:0737-5352-32; available at http://vista.cira.colostate.edu/improve/Publications/improve_reports.htm

(accessed 3/19/2008).
7. Trijonis, J.C., Characterization of Natural Background Aerosol Concentrations, Appendix A in *Acidic Deposition: State of the Science and Technology, Report 24, Visibility Existing and Historical Condition – Causes and Effects*, National Acid Precipitation Assessment Program, 1990.
8. Ames, R.B.; Malm, W.C., Recommendations for Natural Condition Deciview Variability: An examination of IMPROVE Data Frequency Distributions, In (CD-ROM) *Proceedings of the AWMA/AGU Specialty Conference on Regional Haze and Global Radiation Balance – Aerosol Measurements and Models: Closure, Reconciliation and Evaluation*, Bend OR, October, 2001.
9. Tombach, I., *Natural Haze Levels Sensitivity -- Assessment of Refinements to Estimates of Natural Conditions*, Report to the Western Governors Association, January 2008, available at http://www.wrapair.org/forums/aamrf/projects/NCB/Haze_Sensitivity_Report-Final.pdf (accessed 3/19/2008).
10. Ryan, P.A. *Review of the U.S. Environmental Protection Agency Default Implementation Guideline for the Regional Haze Rule*, report number 1011119 prepared for EPRI, Palo Alto, CA, 2004.
11. Ryan, P., Lowenthal, D., Kumar, N. Improved Light Extinction Reconstruction in IMPROVE, J. Air & Waste Manage. Assoc. **55**, 1751-1759, 2005.
12. Pitchford, M., Malm, W., Schichtel, B., Kumar, N., Lowenthal, D., Hand, J., Revised Algorithm for Estimating Light Extinction from IMPROVE Particle Speciation Data, J. Air & Waste Manage, Assoc. **57**: 1326 – 1336, 2007.
13. Natural Haze Levels II: Application of the New IMPROVE Algorithm to Natural Species Concentration Estimates, a final report presentation by the Natural Haze Levels II Committee to the RPO Monitoring/Data Analysis Work Group, July 2006, available at http://vista.cira.colostate.edu/improve/Publications/GrayLit/029_NaturalCondII/naturalhazelevelsIIreport.ppt
14. Visibility Information Exchange Web Site (VIEW), accessible at http://vista.cira.colostate.edu/views/Web/IMPROVE/SummaryData.aspx (accessed 3/19/2008).
15. Visibility Information Exchange Web Site (VIEW), accessible at http://vista.cira.colostate.edu/views/web/documents/substitutedata.aspx (accessed 3/19/2008).
16. Visibility Information Exchange Web Site (VIEW), accessible at http://vista.cira.colostate.edu/views/Web/DataRepository/DataRepository.htm (accessed 3/19/2008).
17. Park, R.J. Jacob, D.J., Kumar, N., Yantosca, R.M., Regional Visibility Statistics in the United States: Natural and Transboundary Pollution Influences, and Implications for the Regional Haze Rule, Atmos. Environ 40: 5405-5423, 2006.