# Examination of U.S. Environmental Regulatory Criteria for Ozone from a Statistical Perspective

Lawrence H. Cox

# NRCSE

Technical Report Series

# EXAMINATION OF U.S. ENVIRONMENTAL REGULATORY CRITERIA FOR OZONE FROM A STATISTICAL PERSPECTIVE

Lawrence H. Cox
U.S. Environmental Protection Agency
National Exposure Research Laboratory (MD-75)
Research Triangle Park, NC 27711 USA
COX.LARRY@EPAMAIL.EPA.GOV

## 1. Introduction

There is recent international interest, e.g, by the U.K. Royal Commission on Environmental Pollution, in using statistics to express and evaluate environmental regulatory criteria (Barnett and O'Hagan, 1997; Cox et al. 1999). In the United States, recent review of National Ambient Air Quality Standards led to revision of the U.S. regulatory criteria for ambient ozone and particulate matter. This in turn led to scientific and policy debates resulting in a decision to postpone implementation of the new regulations pending a second review, scheduled during 2000-2002.

Statistical issues include examining the effects of Types I and II errors arising from sampling and measurement error. Our research is directed at developing a methodology for examining the problem of dealing with uncertainty and variation in environmental regulations and compliance criteria. Our current approach is illustrated through statistical analysis of the recent standard, promulgated in 1979, and the revised (1997) standard for ambient ozone, using data collected at ambient ozone monitors in California's San Joaquin Valley during summer 1990. This paper presents preliminary findings based on quantifying measurement error or precision in terms of small-scale spatial and temporal variability, laying the groundwork for future work. The research team consists of the author (at USEPA) and David Caccia, Peter Guttorp, Paul Sampson and Mary Lou Thompson (all at the National Research Center for Statistics and the Environment, University of Washington, Seattle, WA).

## 2. U.S. Ozone Regulations and Compliance Criteria

The quantity of regulatory interest is the concentration of ozone in ambient air, measured in parts per billion (ppb) or million (ppm), over an appropriate averaging time. Measurements are taken at monitoring sites comprising a *monitoring network* within a *pollution control district*. Ozone is measured at each site for each hour of each day and reported as average hourly ozone concentration. A district records an *exceedance* of the recent ozone *1 hour standard* on a daily basis if during any hour of the day one or more monitors in the district records an average hourly concentration in excess of 0.12 ppm ozone; and, is in *non-attainment status* if at any monitor the *annual expected number of daily exceedances* exceeds one. Several factors affect this important value, e.g., accounting for missing data in determining "actual" annual exceedances, averaging actual exceedances over three years to obtain annual expected exceedances, and, perhaps most important, data rounding: Although many monitors record ozone to the nearest ppb, the standard is set to two decimal places (ppm), requiring rounding to the nearest 10 ppb. A non-attainment area is required to develop an approved air quality remediation plan aimed at achieving compliance with the standard, typically a lengthy and costly process.

The revised ozone standard, the *8 hour standard*, is based on the daily maximum 8 hour average concentration. Non-attainment results if at any site the average over three years of the annual fourth-highest daily maximum 8 hour average concentration exceeds the regulatory standard of 0.08 ppm. Note this important technical difference: non-attainment is defined in terms of a three-year average level of (near) maximum ozone, in lieu of average (or expected) number of exceedances of the standard. See 40 CFR, Part 50.09 and Appendix H, and Part 50.10 and Appendix I, respectively, for details on the 1 and 8 hour standards, available at website: http://www.access.gpo.gov/nara/cfr/.

Data for our study were obtained from the SARMAP study involving 131 ozone sites near the San Joaquin Valley for 62 days during summer 1990, including five ozone episodes. 74 sites were California Air Resources Board (CARB) sites which reported only to the nearest 10 ppb.

## 3. Temporal Assessment of Precision at Monitoring Sites

Available from SARMAP are site-based estimates of *precisions* (instrument or measurement error) for the hourly ozone measurements. We carried out empirical assessments of measurement error from a time series perspective for comparison with reported precisions for the 1 hour standard.

We represent hourly ozone observations at a monitoring site as $y_t = \mu_{h(t)} + x(t) + \varepsilon(t)$, where h(t) is hour of the day at time t, $\mu_{h(t)}$ is the mean concentration for hour of the day h(t), $\mu_{h(t)} + x(t)$ is ``true'' concentration at time t, and $\varepsilon(t)$ is combined measurement error and variations in concentrations at time scales shorter than one hour. Thus, x(t) is assumed to have a temporal variogram continuous at the origin, while $\varepsilon(t)$ represents a *nugget effect*, whose variance we compared with reported precisions. Previous work indicates that time series of hourly ozone concentrations are best modeled on a square root scale, viz., $y(t) = \sqrt{O_3(t)}$. Even after detrending for hourly means, y(t) exhibited periodic autocorrelation. Consequently, we considered only observations within a seven hour window of each daily maximum (which varied between 9 am and 6 pm), which captured the daily maximum while allowing us to assume stationary autocorrelation for the resulting time series (viz., lags up to six hours).

For each site we computed empirical *temporal variograms* of the hourly detrended data, $z(t) = y(t) - \hat{\mu}_{h(t)}$, for lags up to 6 hours: Var(z(t) - z(t + j)), j =1, 2,...,6. We judged that Gaussian variograms adequately characterized these plots. For comparison with reported precisions we transform these values back to ozone (ppb). Assuming that at time t, the observed square-root-transformed ozone concentration is normally distributed about the ``true'' square-root concentration with variance $\sigma^2$, then: $Var(\,O_3(t)) = Var(\,y^2(t)) = 2\sigma^4 + 4\sigma^2 (\,\mu_{h(t)} + x(t)\,)^2$.

Almost all fitted (square-root) nuggets ranged from zero to 0.24, with most in the range 0.03 to 0.16; some were estimated as zero. For true ozone near 120 ppb, standard errors were mostly 2-3 ppb at most non-CARB sites and 4-5 ppb for CARB sites, the latter value consistent with rounding. These values agree well with reported precisions. These values create no loss of precision for deciding whether a site is in exceedance for a particular hour. If the true site/hour ozone is 120 ppb, then the probabilities of declaring the site/hour in violation or in compliance are both approximately 0.5. For true ozone slightly less than 120, there is a small probability of falsely declaring a violation, and an equally small probability of falsely declaring compliance if true ozone is slightly more than 120. These differences are not meaningful and, to two decimal places (ppm), not even detectable. It is important to emphasize, however, that this early work is focused on exceedance at a site at a point in time, and not (yet) on noncompliance, which is more complex and based on three years of data: This research does not (yet) represent an analysis of potential regulatory impacts of accounting for uncertainty and variability in a formal statistical manner, or of not doing so.

## 4. Spatio-Temporal Assessment of Precision at Monitoring Sites

The 1 hour and 8 hour U.S. ozone standards do not address issues of spatial estimation or spatial sampling statistically. Because regulation is done on a regional (pollution control district) basis, it is appropriate to examine the effects of precision of ozone measurements at monitoring sites in terms of its effect on how well the sites characterize the region they are intended to represent. For this purpose we consider a spatio-temporal model for the daily ozone summaries, maximum 1 hour average concentration and maximum 8 hour average concentration: $y(s,t) = \mu(s) + x(s,t) + \varepsilon(s,t)$, where $\mu(s)$ represents long-term daily mean daily at location s, $\mu(s) + x(s,t)$ represents ``true'' ozone concentration for day t at location s, assumed to vary smoothly in time and space, and $\varepsilon(s,t)$ represents combined error due to measurement error and variation in ozone concentrations at spatial scales smaller than the smallest inter-site distance (perhaps a few km). We compute a spatial variogram, or *spatial dispersion function* (Guttorp and Sampson 1994) as: D(s,u) = Var (y(s,t) - y(u,t)). Empirical estimates of this variance are computed from the (up to 62) observations at locations s and u. Sampson and Guttorp model spatial dispersion using nonlinear deformations of the geographic coordinate system in order to account for and model its nonstationary structure. As we are interested only in the nugget effect corresponding to $\varepsilon(s,t)$, determined by D(s, u) for small | s - u| (e.g., a few km to ten's of km), we need only examine simple variogram or dispersion plots of $D(s_i, s_j)$ vs. $| s_i - s_j |$.

For daily maximum 1 hour concentrations, statistical analysis is most appropriate on a logarithmic scale. Robust estimation revealed that the nugget (intercept) of the dispersion scatter is about 0.01 on a log scale, corresponding to a spatial precision of about 10%, or 12 ppb when true ozone concentration is 120 ppb. For daily maximum 8 hour concentration, statistical modeling is most appropriate on a square root scale. The spatial nugget on a square root scale is about 0.1, representing a standard error of 5.7 ppb for an assumed ``true'' 8 hour ozone level of 80 ppb. For comparison with the 1 hour standard, the estimated standard error assuming "true" 8 hour ozone at 120 ppb is 6.9 ppb.

We examined the 1 hour and 8 hour standards of 0.12 ppm and 0.08 ppm via formal tests of hypothesis. Critical values were computed assuming normality for the log and square root transformations of the 1 hour and 8 hour measurements, respectively, converting back to the original scale of measurement and rounding to the nearest 10 ppb. We consider two scenarios, one aiming to protect against false designation of compliance with a standard, and the other aiming to protect against false designation of noncompliance. We use only two months of 1990 data in this preliminary study, not a full three years as for the standard.

Under the first scenario we assume that exceedance of these standards has health effects that must be protected against and consider a testing procedure against the alternative that the standard exceeds the true concentration level, tested at the .10 significance level. Calculations assuming the log and square-root scale precisions result in estimated critical values of 106 ppb for the null hypothesis based on the 120 ppb maximum 1 hour ozone concentration and 73 ppb for the null hypothesis of an 80 ppb maximum 8 hour average ozone concentration. These are rounded to 0.11 ppm for the 1 hour standard of 0.12 ppm and 0.07 ppm for the 8 hour standard of 0.08 ppm. Under the second scenario (protecting against false designation of exceedance) we consider a testing procedure against the alternative that the standard is exceeded by the true concentration, tested at the .10 significance level. Critical values based on the precisions given above for the null hypotheses of 120 ppb and 80 ppb under the 1 hour and 8 hour standards are 136 ppb and 87 ppb, respectively. These result in critical values of 0.14 ppm for the 1 hour standard of 0.12 and 0.09 ppm for the 8 hour standard of 0.08. Test results are summarized in Table 1. Based on these results, we concluded that the (im)precision in the measurement of air quality due to spatial variability can affect the definition or interpretation of regulatory compliance when viewed from the statistical perspective of tests of hypotheses.

*Table 1: Results of Hypothesis Tests:  Number of Sites in Exceedance and Number of Exceedances*

|  | 1 hour standard | 8 hour standard |
|---|---|---|
| Protect against false compliance <br> *A*: Standard > Concentration   $\alpha = .10$ | 90 *sites* <br> 676 *exceedances* | 107 <br> 1814 |
| Current:  Do not account for variability <br> Critical value = Standard         $\alpha = .50$ | 48 <br> 198 | 99 <br> 1244 |
| Protect against false noncompliance <br> *A*: Standard < Concentration   $\alpha = .10$ | 26 <br> 69 | 87 <br> 813 |

## 5.  Acknowledgement

## REFERENCES

Barnett, V. and A. O'Hagan (1997), **Setting Environmental Standards, the Statistical Approach to Handling Uncertainty and Variation**, London: Chapman and Hall.

Cox, L., P. Guttorp, P. Sampson, D. Caccia and M.L. Thompson (1999), "A Preliminary Statistical Examination of the Effects of Uncertainty and Variability on Environmental Regulatory Criteria for Ozone," in: **Novartis Foundation Symposium 220: Environmental Statistics**, London: John Wiley and Sons, Ltd., in press.

Guttorp, P. and P. Sampson (1994), "Methods for Estimating Heterogeneous Spatial Covariance Functions with Environmental Applications," in: Patil, G.P. and C.R. Rao (eds), **Handbook of Statistics XII: Environmental Statistics**, New York: Elsevier/North Holland, 663-690.

## RÉSUMÉ

*Les effects d'incertitude et variabilité sur des decisions et normes régulatrices écologiques sont examinées.*