A Preliminary Statistical Examination of the Effects of Uncertainty and Variability on Environmental Regulatory Criteria for Ozone

Lawrence H. Cox   Peter Guttorp   Paul D. Sampson

David C. Caccia   Mary Lou Thompson

NRCSE

Technical Report Series

NRCSE-TRS No. 016

A PRELIMINARY STATISTICAL EXAMINATION OF THE EFFECTS OF UNCERTAINTY AND VARIABILITY ON ENVIRONMENTAL REGULATORY CRITERIA FOR OZONE

Lawrence H. Cox[1], Peter Guttorp[2], Paul D. Sampson[2], David C. Caccia[2],

and Mary Lou Thompson[2]

[1] U.S. Environmental Protection Agency

National Exposure Research Laboratory (MD-75)

Research Triangle Park, NC 27711

USA

[2] National Research Center for Statistics and the Environment

University of Washington

Box 351720

Seattle, WA 98195-1720

USA

Short title: EFFECTS OF VARIABILITY ON OZONE REGULATIONS

1

A PRELIMINARY STATISTICAL EXAMINATION OF EFFECTS OF UNCERTAINTY AND VARIABILITY ON ENVIRONMENTAL REGULATORY CRITERIA FOR OZONE

Lawrence H. Cox, Peter Guttorp, Paul D. Sampson, David C. Caccia, and Mary Lou Thompson

Abstract

Basing the quantitative expression of environmental regulatory standards and associated compliance criteria on statistical principles has recently received attention in Europe, most visibly a study (Barnett and O'Hagan, 1997) by the U.K. Royal Commission on Environmental Pollution. These issues are timely for consideration in the United States where a recent periodic review of National Ambient Air Quality Standards (NAAQS) has led to revision of the regulatory standards for ambient ozone and particulate matter. Salient statistical issues include accounting for errors of the first and second kind due to sampling and measurement error. These issues appear routine statistically and also may seem absent from regulations, but neither is necessarily the case. This paper is directed towards developing a methodology for examining the problem of dealing with uncertainty and variation in environmental regulations and compliance criteria. Our approach is illustrated through statistical analysis of the (old) 1 hour and the (new) 8 hour standards for ambient ozone, based on intensive monitoring in California's San Joaquin Valley during summer 1990 performed under the SARMAP Project. 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.

Keywords: NAAQS, Air Quality, Measurement Error, Spatial Variability.

1 Introduction

National environmental protection authorities are charged with protecting environmental conditions, including air quality, water quality, and ecosystem health and sustainability. A crucial first step towards environmental protection is to quantify and measure environmental conditions. For the case of air quality, the quantity of interest is often the measured concentration of a pollutant in ambient air, e.g., parts per billion (ppb) ozone. Measurements are obtained through deployment of a network of appropriate monitoring devices.

A subsequent step in environmental protection is regulation. This requires distinguishing between pollutant levels that will and will not be tolerated. In the United States, national ambient air quality standards (NAAQS) in the form of air quality regulations have been in place since the 1970s for each of six *criteria (air) pollutants*: ozone, carbon monoxide, sulphur dioxide, nitrogen dioxide, lead, and particulate matter. These are referred to as criteria pollutants because their standards were set on the basis of health (science) based criteria.

In the United States, air pollution is monitored at (multiple) monitoring sites within predefined pollution control districts. Ozone is measured at each site for each hour of each day and reported as average hourly ozone concentration , typically in parts per billion (ppb) ozone in ambient air. A district records an *exceedance* of the recent 1hour ozone standard on a daily basis if one or more monitors in the region records an average hourly concentration in excess of 0.12 ppm ozone during any hour of the day; and, *fails to meet the standard* if at any monitor in the district the annual *expected number* of daily exceedances exceeds one. Several technical factors comprised this definition, viz., the annual expected number of exceedances at a monitoring site was defined to be the three-year average of *estimated numbers of (annual) exceedances* at the site, which in turn was based on observed

exceedances, adjusted if necessary for missing data. Perhaps most important, although many monitors record ozone to the nearest ppb, the *1 hour ozone standard* was set to two decimal places (ppm), thereby requiring that readings recorded in ppb be rounded to the nearest 10 ppb prior to comparison to the standard. As discussed below, the U.S. EPA has recently proposed a revised standard, based on 8-hour average concentration, which we refer to as the *8 hour ozone standard* (U.S. Federal Register, 1997). Technical details on the 1 hour and 8 hour standards, and their associated compliance criteria, can be found in 40 CFR, Part 50.09 and Appendix H, and 40 CRF, Part 50.10 and Appendix I, respectively (CFR stands for U.S. Code of Federal Regulations; website: /www.access.gpo.gov/nara/cfr/).

National regulations require that an area in *non-attainment status* relative to the standard develop and implement an approved air quality remediation plan aimed at achieving compliance with the standard. This can be a prolonged, complex and expensive process. Regions in non-attainment under the 1 hour standard were given a *classification* ranging from *marginal* to *extreme*, based on a measure of the typical level of exceedance over three years called the *design value*. A design value in excess of 160 ppb classified an area as *serious* or worse, requiring that remediation plans be based on simulations of the effects of anticipated changes in pollutant emissions using complex photochemical models (this volume: Sampson et al., 1999).

In the United States, regulatory standards and associated compliance criteria are reviewed periodically, and on this basis may be continued or revised. This national assessment of air quality standards is based on a synthesis of available scientific information on exposure, health and ecological effects, and exposure-response-effects relationships. Recently (1996-97), U.S. national assessments were completed for ozone and particulate matter, resulting in revised regulations and compliance criteria. The revised regulatory standard for

ozone is 0.08 ppm, based on the maximum of 24 possible 8-hour averages during a day. Compliance with the 8 hour standard is determined as follows: compliance is achieved if at each site the three-year average of the annual fourth-highest daily 8 hour maxima does not exceed the standard. As with the 1 hour standard, computations are to two decimal places. Thus, for a monitor reporting in ppb, compliance is achieved if the computed average is less than 85 ppb. Simulations reveal that the 8 hour standard is stricter than the 1 hour, viz., an EPA analysis (U.S. Federal Register, 1997) based on then-recent data revealed that counties comprising an estimated 133 million people could fail to meet the 8 hour standard, versus 75 million under the 1 hour standard.

In the United States, statistical issues in setting air quality standards were presented at a 1996 workshop cosponsored by the U.S. EPA and the National Institute of Statistical Sciences. Member nations of the European Community are demonstrating interest in statistical approaches to defining environmental regulations and associated compliance criteria. Notably, the U.K. Royal Commission on Environmental Pollution recently commissioned a study on statistical approaches to handling uncertainty and variation in setting environmental regulations, conducted by University of Nottingham Professors Vic Barnett and Anthony O'Hagan and resulting in their book (Barnett and O'Hagan, 1997). The study appears motivated by concern that, although errors (such as due to imprecise environmental measurement) and uncertainty (such as due to sampling variability) evidently affect the accuracy and precision of quantities crucial to regulatory decision making, these factors appear to be unaccounted for in associated compliance criteria.

For example, Barnett and O'Hagan (1997) consider a site at the regulatory boundary of one expected exceedance per year based on the 1 hour standard. Under the assumption of a Poisson model for the number of exceedances in a given year, the probability of more than

three exceedances in a 3-year period (i.e., of violating the standard) when the site is (just barely) in compliance is about 0.35, whence the probability of three or fewer exceedances in a 3-year period for a site (just) in violation of the standard is 0.65 (Barnett and O'Hagan, 1997, pp. 80–81). These computations, while based on assumptions not entirely realistic (viz., that the monitor is at the regulatory threshold on every day), do illustrate potential issues and motivate further consideration.

The purpose of our research is to extend this line of reasoning and explore the feasibility, advantages and disadvantages of statistically-based environmental regulatory standards and compliance criteria. One statistical foundation on which this might rest is hypothesis testing, and the preliminary analysis presented here is based on calculations within a hypothesis testing framework. The research is currently focused on ambient ozone and the 1 hour and 8 hour ozone standards, and some alternatives. Future examination of particulate matter standards is planned. Data on which the calculations are to be based were collected during summer 1990 at approximately 131 sites in central California under the SARMAP project. Future calculations will include data collected at many of these sites over longer time periods.

The research has just begun, and in this paper we present only preliminary concepts and results. The SARMAP project is described in Section 2, and the summer 1990 data used in the hypothesis testing calculations are summarized. Section 3 presents our approach to assessing measurement precision from temporal and spatial perspectives using variograms. These precisions are then used in considering the standards from the perspective of hypothesis testing calculations in Section 4; in particular, the 1 hour and 8 hour ozone standards are compared with each other and with similar standards set (a) to protect against false designation of compliance (health effects receiving foremost consideration) and (b) to pro-

tect against false designation of noncompliance. Plans for future research, including issues of spatial sampling and averaging, appear in Section 5.

2 The SARMAP Study and Database

We use hourly surface ozone data from the San Joaquin Valley and surrounding regions of central California collected in the summer of 1990 as part of the SARMAP project. SARMAP stands for SJVAQS/AUSPEX Regional Model Adaptation Project; where SJVAQS stands for San Joaquin Valley Air Quality Study; and AUSPEX for Atmospheric Utility Signatures, Predictions, and Experiments. The study, which combined several model building and data collection efforts, aimed to develop a comprehensive model addressing ozone, aerosol, visibility, and acid deposition issues and to obtain a high-quality data base for application and evaluation of a complex photochemical model. The goals of the SARMAP study included (Solomon and Silver, 1994):

- Providing decision-makers with an improved understanding of the causes of the ozone concentrations observed in the San Joaquin Valley.

- Providing a capability for comprehensive state-of-the-art Eulerian modeling of atmospheric ozone and secondary aerosol concentration, visibility reduction, and wet and dry acidic deposition in a complex terrain setting.

- Providing field data for operational and diagnostic model evaluation and emission sensitivity analyses.

The SARMAP study assembled aircraft and surface records of many pollutant and/or meteorological variables. Surface measurements were made at 326 monitoring stations in the San Joaquin Valley and environs for 62 days (July 1 through August 31) during the

7

summer of 1990, with intensive sampling during five ozone episodes. Of the surface sites, 131 measured ozone. Because the data collection merged data from several different monitoring networks, the measurements were made on different instruments with different quality assurance requirements. Detailed documentation is available in Solomon and Silver (1994) and references cited therein. Most of the ozone monitoring sites (74 of 131) were under the jurisdiction of the California Air Resources Board (CARB). These sites reported hourly ozone concentrations only to the nearest 10 ppb (consistent with the standard) while most other contractors reported concentrations to the nearest ppb. (Perusal of EPA AIRSWeb reports, available of the internet at www.epa.org/airsweb/monrank.htm, suggests that the CARB have been reporting to the nearest 1 ppb since 1994.) Depending on whether the data are being used in scientific or regulatory analysis, this rounding may account for one source of *measurement error*. It is convenient to present our results in ppb (rather than ppm), but this is done with the understanding that computations can be made in ppb, but comparison to the standard involving ppb must be made at two (and not three) decimal places on a ppm scale.

Figure 1 illustrates the locations of the 131 ozone monitors in the SARMAP database with contours representing kriged interpolations of the two month mean daily maximum 1 hour ozone concentration recorded at these sites (mean taken over the two months of the study). It can be seen that there are some sites whose typical (mean) maximum ozone concentration exceeded 120 ppb. These sites occur near Fresno in the central valley and near Bakersfield toward the southeast limit of the valley where average daily maximum ozone concentrations exceeded 140 ppb. The five highest ozone concentrations recorded in the 8 counties comprising the San Joaquin Valley air basin were: 169, 160, 160, 160, and 160.

Figure 2 depicts the mean daily maximum 8 hour ozone concentration. The 8 hour contours parallel the 1 hour contours quite well. But now the regions with two-month mean 8 hour concentrations in exceedance of the new standard of 0.08 ppm are clearly larger. The five highest 8 hour concentrations recorded in the valley were: 137.5, 134.9, 131.9, 127.5, and 127.5.

Figure 3 depicts the exceedances of the 1 hour and 8 hour standards in the San Joaquin Valley over the two-month study period. Although photochemical modeling efforts addressed only one high ozone episode, Aug 3-6, at least one site is in exceedance on 32 out of 62 days under the 1 hour standard, and 46 out of 62 days under the 8 hour standard.

Figure 4 depicts the site-specific exceedance rates, expressed as the number of exceedances per day (accounting for missing observations). The height and width of the rectangles plotted are proportional to the exceedance rates under the 1 hour and 8 hour standards, respectively. The consistent shape of the rectangles shows that the exceedance rate is higher everywhere under the 8 hour standard.

3 Empirical Assessment of Precision of Hourly Ozone Measurements

The 1 hour standard is defined in terms of the hourly ozone levels at monitoring sites without explicit regard for uncertainty due to measurement error and/or sampling variation. We address first the issue of measurement error at a site from a temporal or time series perspective of the sequences of measurements. We then introduce a notion of precision from a spatial perspective by considering the representativeness of monitoring sites for spatial scales of a few kilometers. We leave for future work the important issues of monitoring

network design or spatial sampling for characterization of a region and how these might be incorporated in the specification of air quality standards and compliance criteria.

3.1 Temporally-based assessment of precision at a monitoring site

The contractors responsible for the field monitoring at all the ozone sites incorporated in the SARMAP database carried out QA analyses based on EPA guidelines that furnish *precisions* for the hourly ozone measurements. We carried out empirical assessments of measurement error from a simple time series perspective for comparison with reported precisions and for selection of measures of variation most relevant for statistical interpretation of the 1 hour standard in a hypothesis testing framework.

At a monitoring site we suppose that the sequence of hourly ozone observations may be represented as

$$y(t) = \mu_{h(t)} + x(t) + e(t)$$

where $h(t)$ is the hour of the day corresponding to time t, $\mu_{h(t)}$ represents the long-term mean ozone concentration for hour of the day $h(t)$, $\mu_{h(t)} + x(t)$ represents the "true" ozone concentration for time t, assumed to vary smoothly as a function of time, and $e(t)$ represents measurement error and variations in ozone concentrations at time scales shorter than the hourly reporting period. In terms of a temporal variogram (see, for example, Diggle, Liang, and Zeger, 1994) this means that the process $x(t)$ is assumed to have a variogram continuous at the origin, while $e(t)$ represents (in geostatistical parlance) a pure *nugget effect*, the variance of which should be compared with contractor reported precisions.

Previous analyses have demonstrated that ozone concentrations are most appropriately considered in terms of such a time series model on a square root scale (Guttorp et al., 1994; Carroll et al., 1997); i.e., we take $y(t) = \sqrt{O_3(t)}$. The temporal structure of $y(t)$ is not quite

10

stationary, but rather demonstrates periodic autocorrelation (Bloomfield et al., 1994), even after detrending for variation in hourly means. Because we are only interested in the hourly concentrations in the temporal neighborhood of the daily maximum, we considered only those observations within a 7 hour window of each daily maximum (which varied between 9 am and 6 pm), believing that this subsample of the time series could reasonably be considered to have a stationary autocorrelation structure for temporal lags up to 6 hours.

Thus for each site we computed empirical variograms of the hourly detrended data, $z(t) = y(t) - \hat{\mu}_{h(t)}$ for lags up to 6 hours from these subsampled time series. As we have data for two months, the numbers of pairs of observations contributing to the empirical variances comprising a variogram, $Var(z(t) - z(t+j))$, $j = 1, 2, ..., 6$, vary from about $n = 360$ ($j = 1$) to $n = 60$ ($j = 6$). We judged that Gaussian variograms adequately characterized these plots and provided the most reasonable basis for an extrapolation to lag $j = 0$, representing the nugget effect on a square root scale.

For comparison with contractor reported precisions we transform these values back to ozone concentrations in ppb. Assuming that at a particular time t, the observed square-root-transformed ozone concentration is normally distributed about the "true" square-root concentration $\mu_{h(t)} + x(t)$ with variance σ^2, determined from the nugget of the fitted variogram, it is easy to show that

$$Var(O_3(t)) = Var(y^2(t)) = 2\sigma^4 + 4\sigma^2(\mu_{h(t)} + x(t))^2.$$

Almost all fitted (square-root scale) nuggets range from zero to 0.24, with most in the range 0.03 to 0.16. (Occasionally, a meaningful nugget could not be estimated as other than zero by the fit of the Gaussian variogram.) For true ozone levels near 120 ppb, these correspond to errors mostly in the range 2-3 ppb for most non-CARB sites and up to 4-5 ppb for the CARB sites, the latter being a value consistent with the rounding of recorded

values to the nearest 10 ppb. These values agree well with contractor reported precisions (Solomon and Thuillier, 1995).

Except for the rounding at CARB sites, these values are sufficiently precise for making decisions as to whether a particular site is in violation for a particular hour. If the true ozone level at a site at a particular hour is 120 ppb, then the probabilities of declaring the site/hour in violation or in compliance are both approximately 0.5. However, for true levels more than a few ppb below 120, there is relatively small probability of falsely declaring a violation, and of course an equally small probability of falsely declaring compliance if the true ozone concentration is more than a few ppb above 120. These differences are not meaningful and, to two decimal places (ppm), are not even detectable. Note that the rounding to two decimal places (ppm) that was carried out in reports from CARB sites until 1994 consistent with the compliance criterion dominates the apparent measurement error. (See summaries of data in the Aerometric Information Retrieval System (AIRS) from the "AIRSWeb" website, /www.epa.gov/airsweb/, for indications of rounding in data reporting.) It is also important to note that we are looking here at exceedance at a particular site at a particular point in time, and not at noncompliance, which is more complicated and normally based on three years of data.

3.2   Spatio-temporally-based assessment of precision of monitoring sites

The 1 hour and 8 hour U.S. ozone standards do not address issues of spatial estimation and address spatial sampling only in terms of guidelines that are not at all statistical. The EPA perspective on siting of ozone monitors is presented in the guidelines for the Photochemical Assessment Monitoring Stations (PAMS) program (/www.epa.gov/oar/oaqps/pams/general.html). In brief, monitors are supposed to be sited to characterize each of four types of areas (the

12

upwind/background region, the maximum emission region, the maximum ozone region, and an extreme downwind region), with each monitor intended to be "representative" of a region of spatial extent varying from less than 1 km up to 10's of km. These spatial scales are consistent with those of the SARMAP photochemical model for the San Joaquin Valley, which computes predictions on a 12 x 12 km grid covering the modeling domain and three 4 x 4 km grids centered on urban areas. Therefore, it is appropriate to address the "precision" of ozone measurements at monitoring sites from the perspective of how well they 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

$$y(s,t) = \mu(s) + x(s,t) + e(s,t)$$

where $\mu(s)$ represents the long-term mean daily summary at location s, $\mu(s) + x(s,t)$ represents the "true" ozone concentration for time (day) t at location s, assumed to vary smoothly in time and space, and $e(s,t)$ represents measurement error and variations in ozone concentrations at small spatial scales of perhaps a few km (spatial scales typically smaller than the smallest inter-site distances).

In this spatio-temporal context we compute a spatial variogram, or *spatial dispersion function* (Sampson and Guttorp, 1992; Guttorp and Sampson, 1994), as

$$D(s,u) = Var(y(s,t) - y(u,t))$$

where the empirical estimates of this variance are computed from the time series of (up to 62) observations at locations s and u. The methodology of Sampson and Guttorp models this spatial dispersion using nonlinear deformations of the geographic coordinate system in order to account for and model nonstationary structure in this spatial dispersion function.

Because we are interested here only in the nugget effect corresponding to the component $e(s, t)$, which is determined by the behavior of $D(s, u)$ for small $|s - u|$ (e.g., on the order of a few km to ten's of km), we do not consider the complete non-stationary modeling of $D(s, u)$ here but examine simple variogram or dispersion plots of $D(s_i, s_j)$ vs. $|s_i - s_j|$.

For daily maximum 1 hour concentrations we find that statistical analysis is most appropriate on a logarithmic scale. The nugget effect (intercept) of the dispersion or variogram scatter cannot be determined precisely, but robust estimation of it indicates a variance of approximately 0.01 on a log scale, corresponding to a spatial precision (standard deviation) of about 10%, or approximately 12 ppb when the true ozone concentration is 120 ppb. This value is used in the computations leading to the hypothesis tests of the next section involving the 1 hour standard, as follows. Based on a normal distribution and inverting the log transformation, estimated test critical values are computed to three decimal places on a ppm scale, followed by rounding to the nearest 10 ppb, to obtain final critical values.

For daily maximum 8 hour concentrations we find that statistical modeling is most appropriate on a square root scale. The spatial nugget on the 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 precision of the maximum 1 hour concentration, this nugget represents a standard error of about 6.9 ppb at a "true" 8 hour ozone concentration of 80 ppb. The estimated nugget standard deviation at 80 ppb, viz., 5.7 ppb, is used in the computations leading to the hypothesis tests of the next section involving the 8 hour standard, as follows. Based on a normal distribution and inverting the sqrt transformation, estimated test critical values are computed to three decimal places on a ppm scale, followed by rounding to the nearest 10 ppb, to obtain final critical values.

4 Hypothesis tests and air quality standards

For purposes of discussion we assume that health effects based standards of 0.12 ppm and 0.08 ppm under the 1 hour and 8 hour average methods, respectively, are to be tested formally. 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. In this preliminary study we tabulate exceedances of standards only for the two months of the 1990 SARMAP field study; we do not carry out tests in terms of the 3-year compliance criteria.

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 true concentration level exceeds the standard, tested at the .10 significance level. Calculations assuming the log and square-root scale precisions (standard deviations) given above 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. Figure 5 maps the exceedance rates for these critical values for contrast with those mapped in Figure 4.

Under the second scenario (protecting against false designation of exceedance) we consider a testing procedure against the alternative that the true concentration is exceeded by the standard, 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. Figure 6 maps

the exceedance rates for these critical values for contrast with those mapped in Figure 4.

The differences between Figures 4, 5, and 6 lead to the conclusion 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. These results, though preliminary and partial, indicate that further investigation would be worthwhile.

5 Future Work

The calculations demonstrated here have been based only on two months of data from one year. They will be extended to assess the interpretation of the standards as actually expressed in terms of three year averages, although the distribution theory must explicitly address the properties of order statistics for the form of the 8 hour standard.

Neither the 1 hour nor the 8 hour ozone standard employs spatial averaging (although the annual PM 2.5 standard does). Future work will look at spatial averaging and use of spatial sampling for monitoring network design relative to defining and assessing effects of variability on air quality standards and compliance criteria. We also intend to move beyond ambient concentrations for a single pollutant to multiple pollutants, viz., ozone and particulate matter, and ultimately would like to consider standards more directly in relation to health effects.

Acknowledgement

or recommendation for use. The opinions expressed are those of the authors; no Agency policy or position should be inferred.

References

[1] Barnett V, O'Hagan A 1997 Setting Environmental Standards, the statistical approach to handling uncertainty and variation. Chapman and Hall, London

[2] Bloomfield P, Hurd H, Lundi R 1994 Periodic correlation in stratospheric ozone data. J Time Series Anal 15:127-150

[3] Carroll RJ, Chen R, Li TH, Newton HJ, Schmiediche H, Wang N, George EI 1997 Trends in ozone exposure in Harris County, Texas. J Amer Statist Assoc 92:392-415 (with discussion)

[4] Diggle PJ, Liang K-Y, Zeger SL 1994 Analysis of Longitudinal Data. Oxford University Press, New York

[5] Guttorp P, Sampson PD 1994 Methods for estimating heterogeneous spatial covariance functions with environmental applications. In: Patil GP, Rao CR (eds) Handbook of Statistics XII: Environmental Statistics, Elsevier/North Holland, New York p 663-690

[6] Guttorp P, Meiring W, Sampson PD 1994 A space-time analysis of ground-level ozone data. Environmetrics 5:241-254

[7] National Research Council 1991 Rethinking the Ozone Problem in Urban and Regional Air Pollution. National Academy Press, Washington DC

[8] Sampson PD, Guttorp P 1992 Nonparametric estimation of nonstationary spatial covariance structure. J Amer Statist Assoc 87:108-119

[9] Solomon PA, Silver TA (eds) 1994 Planning and Managing Air Quality Modeling and Measurement Studies: a perspective through SJVAQS/AUSPEX. Lewis Publishers/Pacific Gas and Electric Company, Chelsea, MI, USA

[10] Solomon PA, Thuillier A 1995 SJVAQS/AUSPEX/SARMAP 1990 Field Measurement Project v 2. Field Measurement Characteristics. PG&E report 009.2-94.1, Dept. of Research and Development, San Ramone, CA

[11] U.S. Federal Register 1997 National Ambient Air Quality Standards for Ozone; Final Rule 62 FR 38856. July 18, 1997

Figure Legends

Figure 1. Mean of daily maximum 1 hour average ozone concentrations for July-August, 1990. Contours represent kriged interpolations of mean ozone concentrations computed for each of the 131 monitoring sites in the SARMAP database.
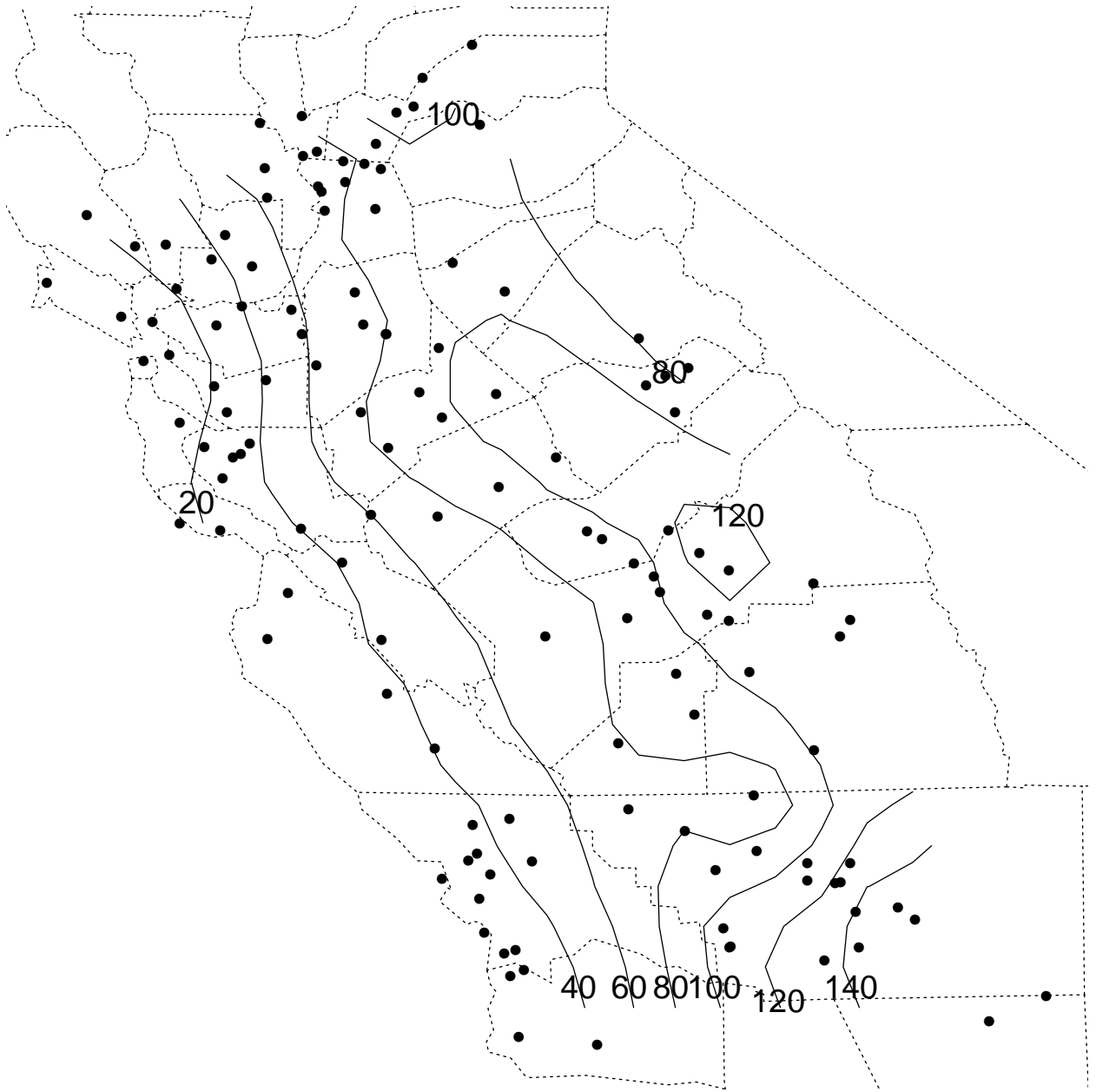
Figure 2. Mean of daily maximum 8 hour average ozone concentrations for July-August, 1990. Contours represent kriged interpolations of mean ozone concentrations computed for each of the 131 monitoring sites in the SARMAP database.

Figure 3. Plot of study day (July-August, 1990) vs. fraction of the 46 monitoring sites in the 8 counties of the San Joaquin Air Basin that were in violation of the 1 hour and 8 hour standards: 0.12 ppm maximum 1 hour ozone concentration, solid line; 0.08 ppm maximum 8 hour average ozone concentration, dashed line.
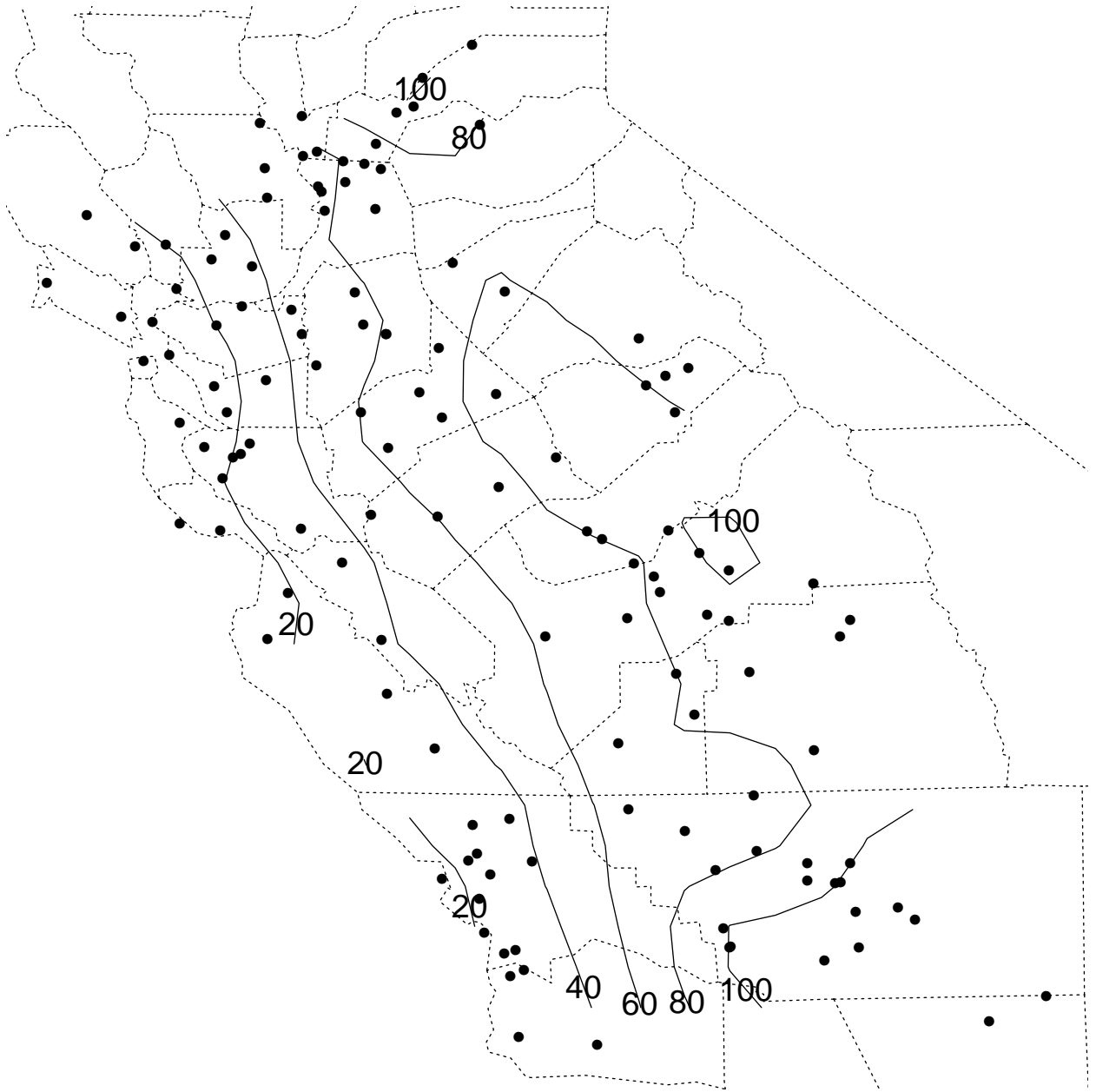
Figure 4. Rectangle map of the exceedance rate expressed as number of exceedances per day (accounting for missing observations) for the 1 hour (0.12 ppm) and 8 hour (0.08 ppm) ozone standards at each monitoring site: height proportional to exceedance rate for 1 hour standard; width proportional to exceedance rate for 8 hour standard. The exceedance rate is consistently higher for the 8 hour standard.
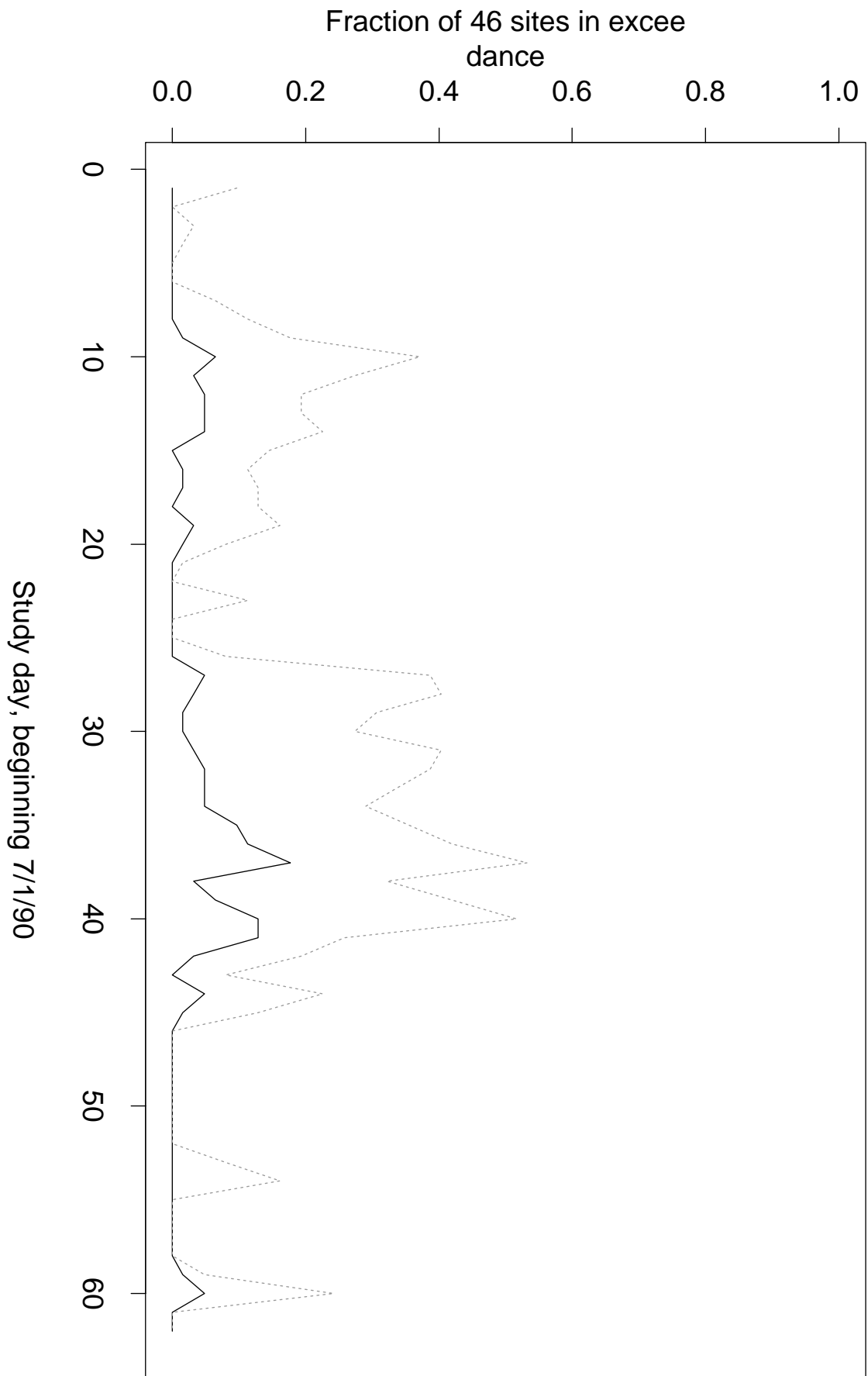
Figure 5. Map of the exceedance rates (as in Figure 4) for standards based on .10-level hypothesis tests of the 1 hour and 8 hour ozone standards against the alternative of false designation of compliance. The critical values defining exceedance are 0.11 ppm for the null hypothesis of a 0.12 ppm max 1 hour ozone concentration and 0.07 ppm for the null hypothesis of an 0.08 ppm max 8 hour average ozone concentration.
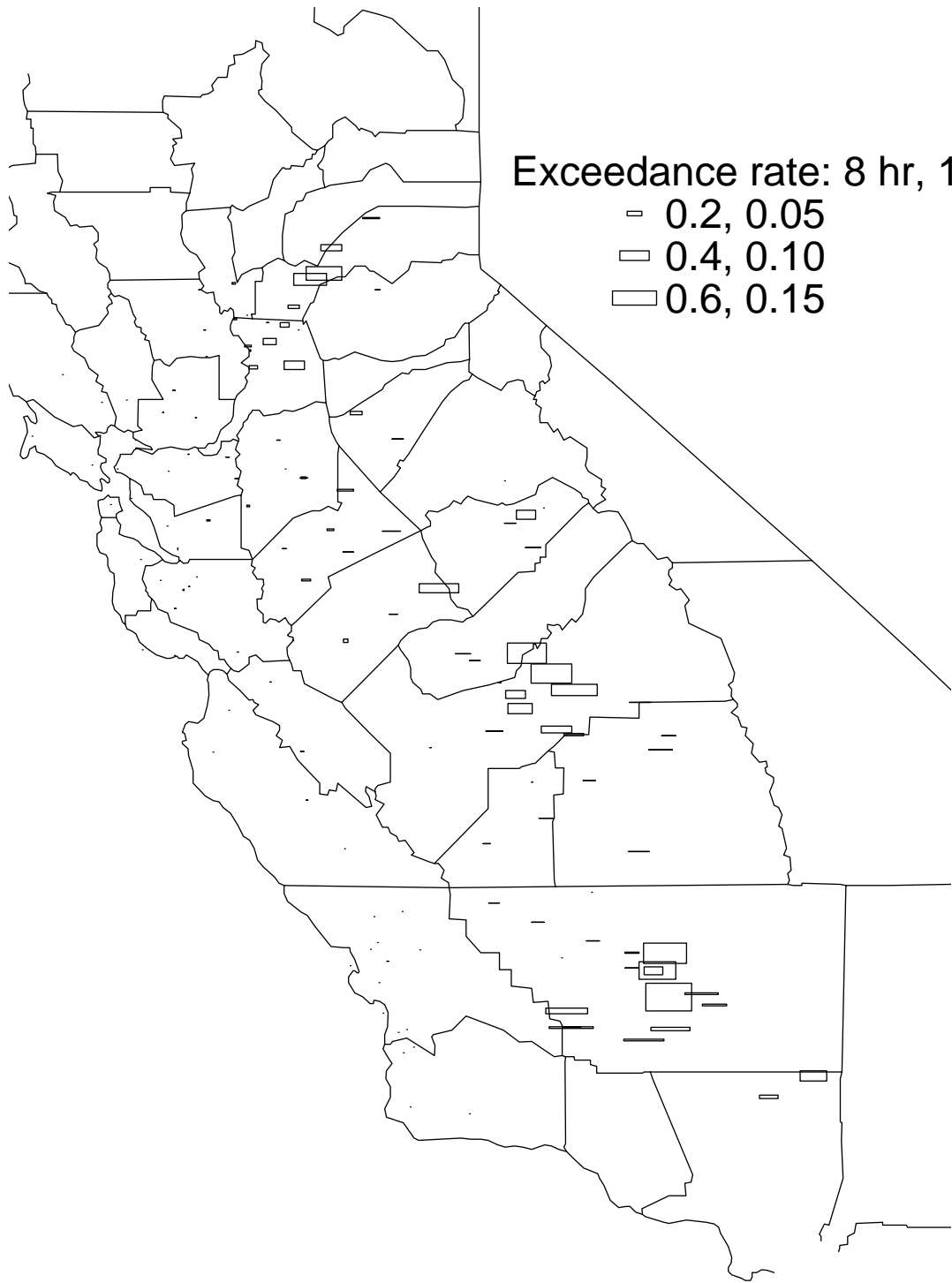
Figure 6. Map of the exceedance rates (as in Figure 4) for standards based on .10-level hypothesis tests of the 1 hour and 8 hour ozone standards against the alternative of false designation of noncompliance. The critical values defining exceedance are 0.14 ppm for the

null hypothesis of a 0.12 ppm max 1 hour ozone concentration and 0.09 ppm for the null

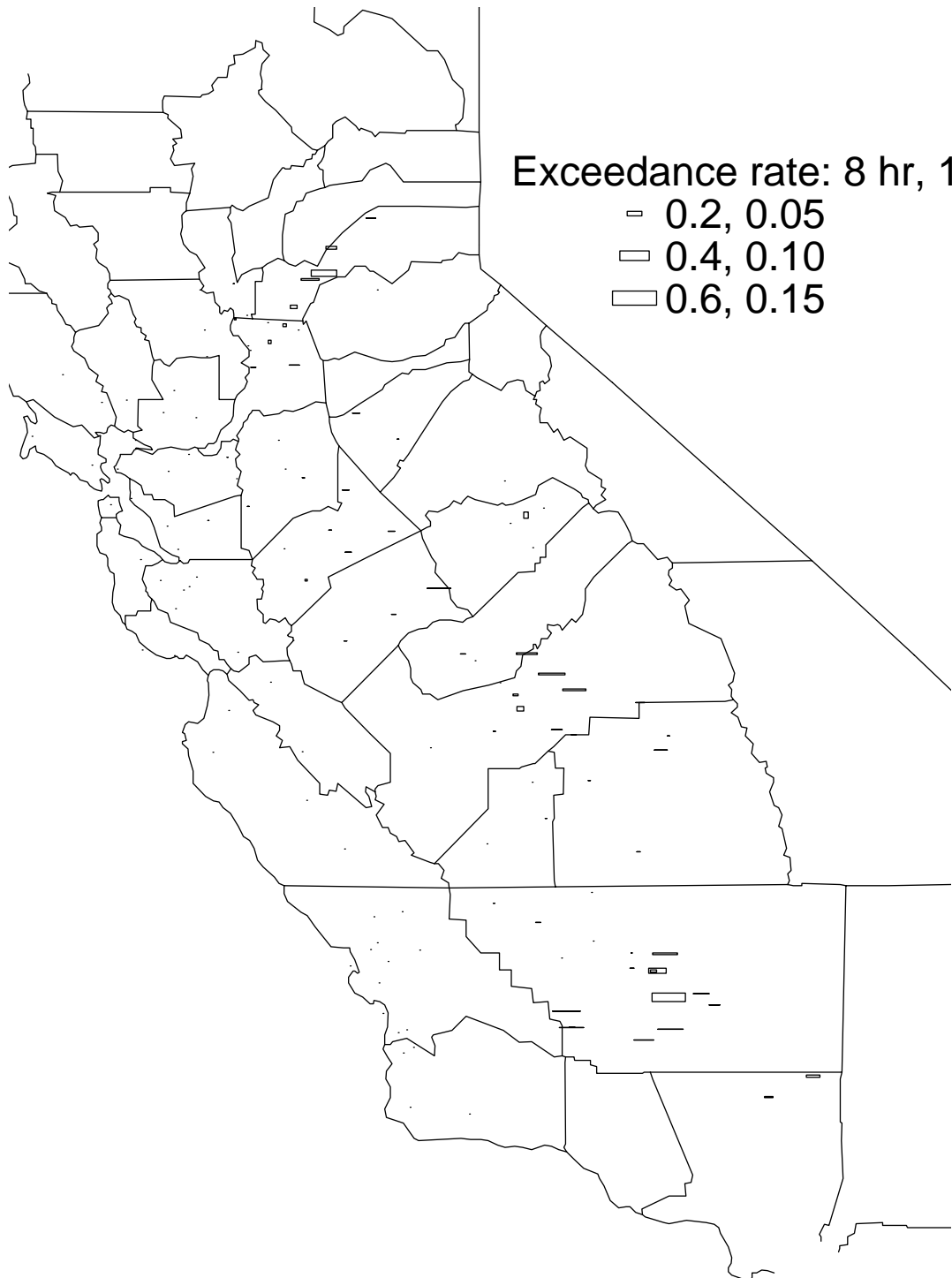hypothesis of an .08 ppm max 8 hour average ozone concentration.

Fraction of 46 sites in excee
dance

Study day, beginning 7/1/90

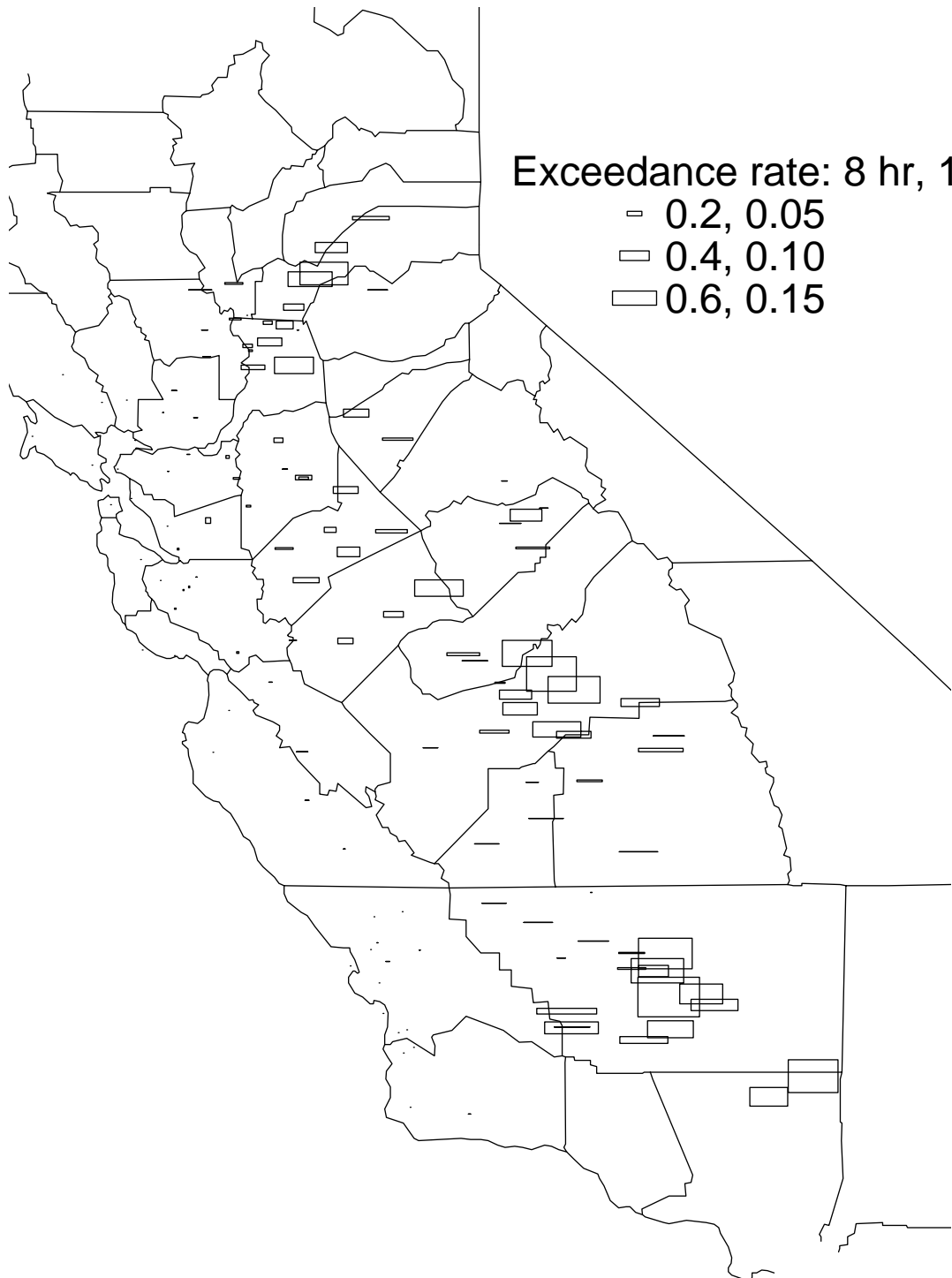Exceedance rate: 8 hr, 1 hr
0.2, 0.05
0.4, 0.10
0.6, 0.15

Exceedance rate: 8 hr, 1 hr

- 0.2, 0.05
- 0.4, 0.10
- 0.6, 0.15

Exceedance rate: 8 hr, 1 hr
- 0.2, 0.05
- 0.4, 0.10
- 0.6, 0.15