

EXAMINATION OF INTER- VARIABLE RELATIONSHIPS AND IMPLICATIONS FOR EPISODE SELECTION USING THE 1999/2000 SEARCH DATABASE

Sharon Douglas
ICF Consulting/SAI
San Francisco/San Rafael, CA

VISTAS Monitoring Workshop
15-16 January 2002
Atlanta, GA



USE OF MODELS TO ADDRESS PM_{2.5} AND REGIONAL HAZE ISSUES

- Theoretical analysis
 - physical and chemical processes
 - nature and extent of problem
- Future-year analysis of effects of emissions changes on air quality, visibility, and health
- Regulatory compliance
 - PM_{2.5} attainment demonstration modeling
 - compliance with regional haze goals (for national parks and wilderness areas)

KEY ISSUES ASSOCIATED WITH PM/REGIONAL HAZE MODELING

- Availability, representativeness, and sufficiency of measured data
- Incomplete understanding of the physical and chemical processes that govern secondary particulate formation
- Uncertainty wrt whether models are able to provide accurate/meaningful results and information

APPROACHES TO PM/REGIONAL HAZE MODELING

- Continental-scale; coarse grid resolution; full-year simulation period
 - REMSAD
- Regional- to urban-scale; finer grid resolution; multi-day simulation periods
 - CMAQ
 - URM (SAMI)
 - UAM-VPM

ISSUES ASSOCIATED WITH PM/ REGIONAL HAZE MODELING APPROACHES

- Full-year simulation period approach
 - representativeness of the modeled year
 - sufficiency of grid resolution (CPU requirements)
 - quality of input fields
- Discrete, multi-day simulation (episode) periods approach
 - representativeness of simulation periods (type, range, frequency of events)
 - use of results from discrete episode periods to estimate annual metrics and health effects

SEARCH EPISODE-SELECTION ANALYSIS OBJECTIVES

- Design, test, and apply a method for selecting discrete episodes for modeling that represent
 - PM_{2.5} and regional-haze events (frequency and magnitude)
 - variations in the PM constituents
 - geographic scales of interest (SE regional, urban)
 - interactions with ozone and other pollutants
 - potential biases in the attainment demonstration procedures

SEARCH EPISODE-SELECTION ANALYSIS OBJECTIVES (CONT)

- Examine whether use of a less intensive database (spatially and temporally) would compromise episode selection
- Suggest improvements to current EPA guidance

SEARCH EPISODE-SELECTION ANALYSIS METHODOLOGY

- Prepare multi-variate database for 1999/2000
 - SEARCH PM, gaseous, and meteorological data
 - regional meteorological data
 - measured variables and constructed parameters
- Apply statistical analysis to summarize and characterize the dataset

SEARCH EPISODE-SELECTION ANALYSIS METHODOLOGY

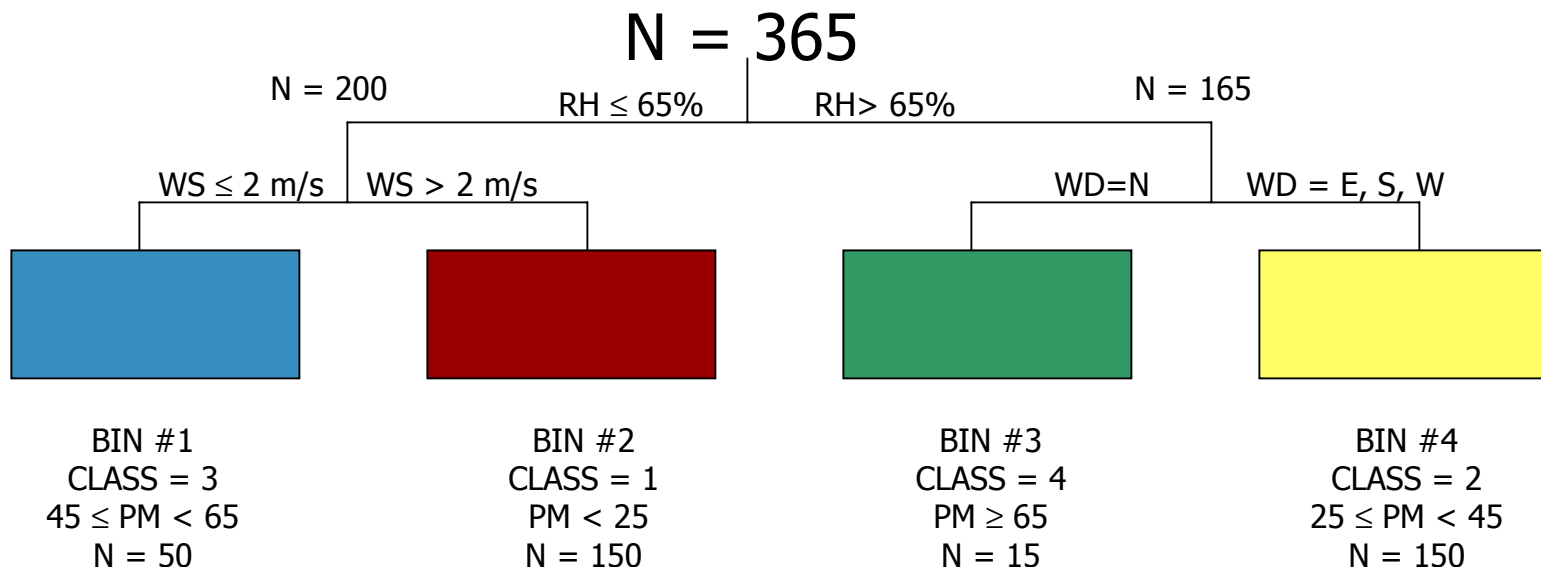
- Use Classification and Regression Tree (CART) analysis to
 - classify days/periods for 1999 according to
 - PM mass
 - PM speciation
 - estimated visibility
 - other dependent variables
 - group days with similar PM/visibility and meteorological characteristics
 - investigate inter-variable relationships
 - determine/infer causal relationship

CLASSIFICATION AND REGRESSION TREE (CART)

- Statistical tool used to classify “objects” (e.g. days) into classification “bins”
 - bins are associated with a certain value or range of a classification variable (e.g., $PM_{2.5}$ concentration)
 - classification is based on the value of other (e.g., meteorological, ozone concentration) parameters
 - for this analysis, will provide information about the conditions that are associated with different ranges of PM concentrations (or visibility) as well as the frequency of different types of “regimes”

CLASSIFICATION AND REGRESSION TREE (CART)

- Output takes the form of an up-side-down classification “tree” - branches/splits and independent variables (data) determine the binning



CART APPLICATION

- Separate analysis for each site
- Use full year of data (1999/2000)
- Classification (dependent) variable options
 - $PM_{2.5}$ (mass)
 - visibility
 - SO_4^-
 - other combinations of variables?

CART APPLICATION

- **Independent variables**
 - PM/visibility parameters for other days/sites
 - ozone concentrations, other gasses?
 - site-specific meteorological parameters
 - upper-air meteorological parameters
 - synoptic-scale meteorological regime

SEARCH EPISODE-SELECTION ANALYSIS METHODOLOGY

- Optimization approach will be applied to identify/select appropriate modeling episode periods
 - using similar methods to those developed/applied for SAMI
 - objective episode selection can accommodate
 - single site
 - single area/multiple sites
 - multiple geographical areas
 - allows one to specify, apply, and evaluate criteria

SPECIFIC OBJECTIVES OF THE EPISODE-SELECTION COMPONENT

- Episode selection analysis will address:
 - development of criteria for episode selection that are supported by the SEARCH data
 - utility of EPA-recommended approaches
 - number of modeling episodes/days; multiple geographical areas
 - modeling for annual vs. 24-hour $PM_{2.5}$ standard
 - modeling for $PM_{2.5}$ vs. regional haze
 - implications of using a more limited dataset