Assignment for STK 4150 fall 2003

Due data: November 28, 2003

November 11, 2003

We will in this assignment analyse data on the 1999 network for PM2.5 (particulate matter of aerodynamic diameter 2.5 microns or less), restricted to the 74 monitors in the states of North Carolina (NC), South Carolina (SC) and Georgia (GA). Much of the interest in this data set arises from the proposed PM2.5 standard, which set a limit of 15 (micrograms per meter cubed) for the long-term mean of PM2.5 in each region.

For data stability reasons, the recorded data are in terms of natural logarithms of mean PM2.5 at each site.

The data are given on different files which are available from the course home page.

pm25A.txt contains the following:

    Col. 1: Longitude of i'th station (i=1,2,...,74)

    Col. 2: Latitude of i'th station

    Col. 3: Estimate logarithm of the mean PM2.5 at i'th station

    Col. 4: Standard error of the quantity in Col. 3

pm25B.txt contains covariates given as indicators of belonging to state (NC in column 1, SC in column 2, GA in column 3) and indicators of land use (agricultural in column 4, commercial in column 5, forest in column 6, industrial in column 7, residential in column 8)

pm25C.txt contains site to site correlations associated with the data in pm25A.txt (a 74×74 matrix).

pm25D.txt contains a trial data set for the kriging locations: longitude, latitude, and the 8 covariates as given in pm25B.txt.

(a). Based on the data above, find an approperiate model.

Use this model to construct two spatial maps: one of the mean PM2.5 at each location in the pm25D.txt data file, and one of the prediction standard error at each location.

Don't forget that the original data are logs of mean PM2.5, but that the maps should be produced at the actual scale of PM2.5, not the log.

(b). Estimate the fraction of the total area of the three states that (as of 1999) is in violation of the 15 microgram per meter cubed standard.