

## Evaluacion

Eligir un juego de datos en Geoestadística y otro en Datos en una red. Tratar los datos y redactar un informe, explicando el problema, los métodos para resolverlo, analizar y comentar los resultados.

El código, estará en un archivo aparte del informe.

El informe se ha de entregar en el obligatorio de la Eva antes del 5 de abril 2019. Por cualquiera pregunta, mandar un e-mail a Liliane.Bel@agroparistech.fr.

### 1. Geoestadística

#### (a) Polución por el ozono en la región de París

Se desea realizar un mapa cotidiano de la concentración en ozono en la región de París. Se dispone de las salidas de un modelo físico determinista elaborado en el Laboratoire de Météorologie Dynamique (Ecole Polytechnique) y de medidas de concentración realizadas por AirParif (la agencia de vigilancia del aire para la región de París) en 21 estaciones.

El archivo *stationsKm.txt* contiene una tabla con las variables (columnas)

- x : abscisas (en km) de las estaciones
- y : ordenadas (en km) de las estaciones
- z : medidas de la concentración máxima del día en las estaciones (en  $\mu/m^3$ )
- mod : valor de la concentración de ozono dada por el modelo físico

El archivo *grilleKm.txt* contiene

- x : abscisas (en km) de los puntos de una grilla que cubre la región
- y : ordenadas (en km) de los puntos de la grilla
- z : valores de concentración dadas por el modelo físico en los puntos de la grilla

Trazando el mapa de la concentración en ozono dada por el modelo físico y comparándolo con los valores observados en las estaciones se puede averiguar que el modelo procura bastante sesgo. En cambio el mapa producido por kriging a partir de las observaciones no tiene sentido físico. Se puede tratar de corregir el mapa del modelo físico haciendo un kriging de los errores medidos en las estaciones.

Se admite la presencia de un ruido de medida de  $10\mu/m^3$  en las observaciones.

#### (b) Olas en el golfo de Lion

El golfo de Lion es un golfo del mar Mediterráneo entre Francia y España. Condiciones climáticas favorables suelen producir olas importantes que pueden ocasionar daños importantes. Para una mejor comprensión del fenómeno se desea producir un mapa de la altura significativa máxima de las olas en todo el golfo. Para ello disponemos de los datos de batimetría (profundidad) en todo el golfo, y de las alturas significativas máxima obtenidas por un simulador numérico en algunos sitios del golfo.

El archivo *Stationsan.txt* contiene para 97 puntos del golfo las coordenadas geográficas (longitud, latitud en grados), los valores del máximo anual de la altura significativa registrada para el año *an* y los valores de batimetría.

El archivo *BathyGL.txt* contiene las coordenadas geográficas de puntos de una grilla fina que recubre el golfo y los valores asociados de batimetría.

Para cada año establecer la relación entre la batimetría y la altura significativa. Corregir el mapa obtenido haciendo un kriging de los residuos.

### 2. Datos en una red

#### (a) Columbus OH spatial analysis data set

Se estudia la criminalidad en la ciudad de Columbus (Ohio, US) en función de variables económicas : ingresos y precio del alojamiento. Los datos Columbus forman parte del paquete R *spdep*. Columbus está dividido en 49 distritos. Para cargar los datos *data(columbus)*. Para ver la descripción de las variables *?columbus*. Para trazar el mapa de los distritos *plot(columbus, col=colores)*,

*colores* es un vector de colores asociados a los distritos. Estudiar la presencia de autocorrelacion espacial en la criminalidad (variable CRIME) en Columbus y proponer un modelo que explica la criminalidad en funcion de las variables economicas.

(b) Cancer de pulmon en Francia

Se estudia la influencia de la actividad industrial en el cancer de pulmon en Francia en los años 60. El archivo *cancer.txt* contiene 8 variables para 82 departamentos :

- poumon : incidencia del cancer
- metallurgie : industria metalurgia
- mecanique : industria mecanica
- textile : industria textil
- vente.tabac : venta de tabaco
- numero.departement : numero del departamento
- latitude : latitud en grados
- longitude : longitud en grados

Para trazar los limites de los departamentos se carga el paquete *maps*, y *map("france")*. Para trazar los departamentos con colores que dependen de una variable se carga la funcion "*trace-France.R*", y *traceFrance(variable,title,color)*. *title* es un titulo y *color* es un vector de colores asociados a los departamentos.

Estudiar la presencia de autocorrelacion espacial en la incidencia del cancer del pulmo y proponer un modelo que explica la incidencia del cancer de pulmon en funcion de las variables de industrias.