

1. The mackerel data

A multi-country survey of mackerel eggs was carried out in the Eastern Atlantic in 1992. The aim was to estimate the total biomass of spawning mackerel. However, a first step is to construct a model which describes the pattern of egg counts in the water samples collected. A map of the sampling region can be produced by

```
sm.script(mackmap)
```

Data from the survey are available through

```
provide.data(mackerel)
```

Produce some initial plots to explore the relationship between the egg density (`Density`) and latitude, longitude, depth, temperature and salinity (`mack.lat`, `mack.long`, `mack.depth`, `Temperature`, `Salinity`). Consider whether any of these variables need to be transformed in order to put them on more suitable scales. (For example, try `Density` on a log scale.)

You may also find it helpful to use `sm.regression(x, y)` to highlight the underlying pattern in these scatterplots.

Consider also the spatial effect by

```
position <- cbind(mack.lat, mack.long)
sm.regression(position, log(Density))
```

Now fit a gam model and use it to identify what the effects of the different explanatory variables are. Instead of the `s` notation used earlier, use `lo` instead, as this will allow terms such as `position` to be entered. For example, you might like to try

```
model <- gam(log(Density) ~ lo(Temperature) +
             lo(log(mack.depth)) +
             lo(Salinity) + lo(position))
par(mfrow = c(2, 2))
plot(model)
par(mfrow = c(1, 1))
```

Use

```
anova(model)
```

to identify whether some of the explanatory variables can be dropped from the model. What is the preferred depth at which mackerel lay their eggs?

2. The Spanish mackerel data

The data collected by Spanish vessels differs from that of most other countries, in that it is largely in the form of presence or absence of eggs in each water sample. These data are available through

```
provide.data(smacker)
```

Use `names(smacker)` to see the variables which are available. Create an indicator of presence or absence by

```
Presence <- as.numeric(Density > 0)
```

Use the `family = binomial` argument to construct a gam for these data, using temperature, depth and spatial position as explanatory variables. What do you conclude about the effects of these covariates?

Is there any difference in the indication of preferred depth at which mackerel lay their eggs?