Data exploration, regression, GLM and GAM course

Highland Statistics Ltd www.highstat.comn

Exercise 2: Data exploration Bailey data

Data description

Since 1979 fisheries surveys have been taken in an area of the NE Atlantic Ocean, assembling a unique fishery-independent dataset from trawls conducted on commercial fishing grounds (800-1500 m), and beyond on the slope (1500-4000 m) and abyssal plain (to 4800 m). Sampling took place in two sampling periods:

1. The "Early" period (1979 to 1989) is before and during the development of the fishery.
2. The "Late" period (1997 to 2002) is considered post commercial fishing.

Gear and techniques used were identical throughout. See Bailey et al. (2008) for details of sampling. The response variables are Dens and TotAbund.

The file Baileyetal2008.xls contains fish abundance data. Each row in the text file represents one site (trawl). The variables are given in Table 1.

Table 1. Variables in the file baileyetal2008.txt.

Variable	Description	Type
TotAbund	Total abundance of all fish at a site (trawl)	Count, response variable
Dens	Density of all fish (= total abundance / sweeping area)	Continuous response variable
MeanDepth	Mean depth of a trawl	Continuous explanatory variable
Year	Year of sampling	Explanatory variable
Period	Time period	Categorical explanatory variable
Xkm and Ykm	Spatial position	Continuous variable
SweptArea	The swept area during a trawl (= sampling effort)	Offset variable

References

- DM Bailey, MA Collins, JDM. Gordon, AF Zuur, IG Priede. (2008) Long-term changes in deep-water fish populations in the North East Atlantic: a deeper-reaching effect of fisheries? Journal Proceedings of the Royal Society: B.

Underlying question and task

The underlying question is whether there is a relationship between density and depth has changed over time. The task for the moment is to apply a data exploration.