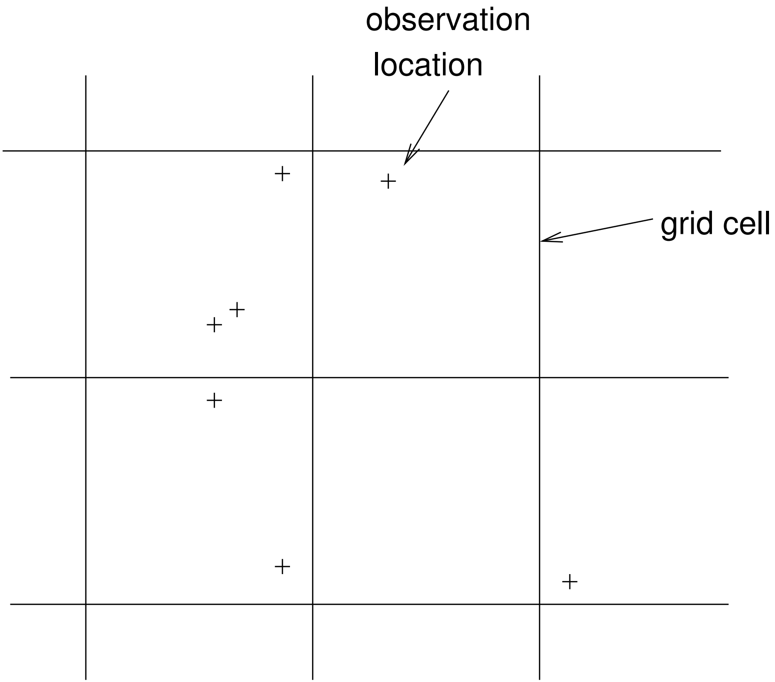
Remote Sensing and Image Analysis

Edzer Pebesma

October 22, 2008

ifgi
Institute for Geoinformatics
University of Münster


observation
location
grid cell

Support

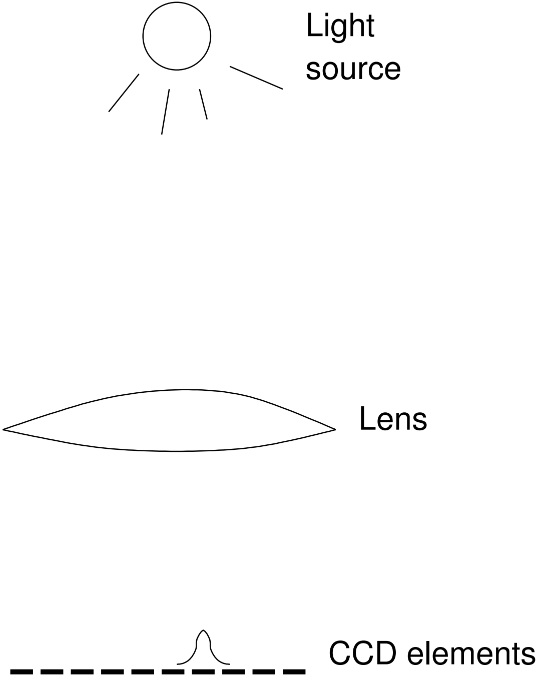
Let $Z(s)$ be the process we're interested in. Let s be a *point* in space with known coordinates: e.g. the location of a sensor. Theoretically (mathematically) a point has no volume. Measurements always refer to matter with some volume (or area). The block average over block B_0 is defined as

$$Z(B_0) = \frac{1}{|B_0|} \int_{B_0} Z(u)du$$

and is usually approximated by a sum of q points s_j that discretize the block B_0 in some way (e.g. regularly)

$$Z(B_0) \approx \frac{1}{q} \sum_{j=1}^{q} Z(s_j)$$

If we have observations on points $Z(s_i)$, and want to estimate the unknown quantity $Z(B_0)$, then we call this a *change of support* problem.


Light
source
Lens
CCD elements

Images and support

Theoretical block averages do not work out completely for imagery data:

- grid cell (pixel) values do represent more of the center of the pixel than of the borders (i.e., non-linear average);
- pixel values contain some information of neighbour pixels (blocks overlap)
- image processing may remove some of the resulting blurr (contrast enhancement)

Estimating block means from points

a.k.a. the problem of up-scaling.

Options are:

- simple mean of observations in a block
  - \+ no model needed
  - \+ estimates are independent
  - \- many pixels will be empty
- block kriging (interpolation)
  - \+ all pixels get an estimate + std.err.
  - \- needs a stationary model for spatial variation
  - \- estimates are no longer independent
  - \- conditional bias due to smoothing

Hard to say ahead of time what will be best in terms of *power*.

Estimating points from block means

Problem of down-scaling; think also sub-pixel classification.
Needs many assumptions of behaviour on the point scale.
May make sense if additional information (roads?) is available:
how do we redistribute a total/average (pixel value) over the grid,
given we know sources etc?

Simple tools

- compare (scatter plot) the in situ sensor values with the block mean (image pixel) values
- try to assess the variability you might expect: variance of points within blocks. One-way ANOVA?
- Is constant variance a reasonable assumption, or need we transform variables (power? sqrt? log?)

Relating two continuous variables

With a straight line: correlation/regression:

$$y_i = a + bx_i + e_i$$

$$y(s_i) = a + bx(s_i) + e(s_i)$$

With a polynomial: regression

$$y(s_i) = a + bx(s_i) + cx^2(s_i) + e(s_i)$$

With a complex function: non-linear regresion, ...

$$y(s_i) = a + bx^c(s_i) + e(s_i)$$

Relating a continuous variables to a categorical variable

ANOVA (Analysis of variance); linear models – linear regression

$$y(s_i) = \mu_j x(s_i) + e(s_i)$$

with j the number of categories.

Relating two categorical variables

χ^2 test for a *contingency table*

	$PM_{10} < 50$		$PM_{10} > 50$	
urban	20	32.4	25	12.6
rural	70	57.6	10	22.4

with O_{ij} and E_{ij} the observed and expected frequencies in cell i, j:

$$\chi^2 = \sum_{i=1}^{p} \sum_{j=1}^{q} \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}}$$