Trend-Surface Analysis

Global versus local spatial interpolation

- Aims of interpolation: establish a model for the spatial variability including any factors affecting it. Prediction of values within the study area at locations where not data have been sampled.
- Global interpolators use all available data to provide predictions for the whole study area of interest.
- Global interpolators generate a smooth surface, which captures small-scale (i.e. large area) global trends related to the first order **expectation level** of a spatial process.
- If any heterogeneity in expected level of the spatial process is not captured properly the data then will show ***second order variability*** in the covariance structure.
- Global interpolators are mostly used to remove the effects of global variation caused by major trends in the surface in order to achieve ***first order stationary in the residuals***.
- Local interpolators are calibrated based on data within small sliding windows around the point at which interpolation is performed. This ensures that estimates are not affected by global heterogeneity but makes the local estimates correlated because they share a subset of the observed data.
- Local interpolators operate either mechanically by ***assuming second order dependency*** through a distance decay function or they explicitly incorporate an ***estimate of the spatial dependence*** at the local level.
- NOTE: if the data are not spatially autocorrelated any local interpolator becomes meaningless.

Trend-Surface Analysis

- Let's denote the longitude of a point *i* by x_i and its latitude by y_i or in vector form by $\mathbf{s}_i = (x_i, y_i)^T$.

- The set of measurements of a dependent variable y_i and perhaps additional covariables $\mathbf{z}_i$ at given coordinates $\mathbf{s}_i$ constitute the geo-reference sample observations.
- Trend surface analysis is a special form of regression analysis defined on point locations. The dependent variable is regressed against a polynomial function of point coordinates (independent variables), i.e., $y_i = f_y(\mathbf{s}_i) + \varepsilon_i$.
- Interpolation of the surface Y at any location of the study region can be conducted by predicting $\hat{y}_i = \hat{f}_y(\mathbf{s}_i)$. Even though these locations do not belong to the sample, we use their coordinates $\mathbf{s}_j^{pred}$ and predict the expected value $\hat{y}_j$.
- Additional co-variables can be added to the equation: $y_i = f_y(\mathbf{s}_i) + \mathbf{z}_i^T \cdot \boldsymbol{\beta} + \varepsilon_i$
- However, interpolation is impossible unless the co-variable surface $\mathbf{z}_j$ is available at all prediction locations $\mathbf{s}_j^{pred}$.

So, Co-variable must be observed, not be predicted.

- While we could estimate the value of the covariates $\mathbf{z}_j$ at any location by yet another trend-surface $\mathbf{z}_i = f_z(\mathbf{s}_i) + \varepsilon_i$, the predicted values $\hat{\mathbf{z}}_i$ are collinear with the coordinates $\mathbf{s}_i$ and, therefore, with the coordinates of $\mathbf{s}_i$ in the trend-surface $f_y(\mathbf{s}_i)$. This perfect multicollinearity makes this approach impossible.
- Trend-surfaces are defined by their polynomial order

Trend- Surface Order	**Linear Function of Coordinates**
0	$\hat{f}(\mathbf{s}_i) = b_{00}$
1	$\hat{f}(\mathbf{s}_i) = b_{00} + b_{10} \cdot x_i^1 + b_{01} \cdot y_i^1$
2	$\hat{f}(\mathbf{s}_i) = b_{00} + b_{10} \cdot x_i^1 + b_{01} \cdot y_i^1 + b_{02} \cdot x_i^2 + b_{11} \cdot (x_i^1 \cdot y_i^1) + b_{02} \cdot y_i^2$
3	$\hat{f}(\mathbf{s}_i) = b_{00} + b_{10} \cdot x_i^1 + b_{01} \cdot y_i^1 + b_{02} \cdot x_i^2 + b_{11} \cdot (x_i^1 \cdot y_i^1) + b_{02} \cdot y_i^2$ $+ b_{30} \cdot x_i^3 + b_{21} \cdot (x_i^2 \cdot y_i^1) + b_{12} \cdot (x_i^1 \cdot y_i^2) + b_{03} \cdot y_i^3$

- Graphs of these orders and their specific representations:

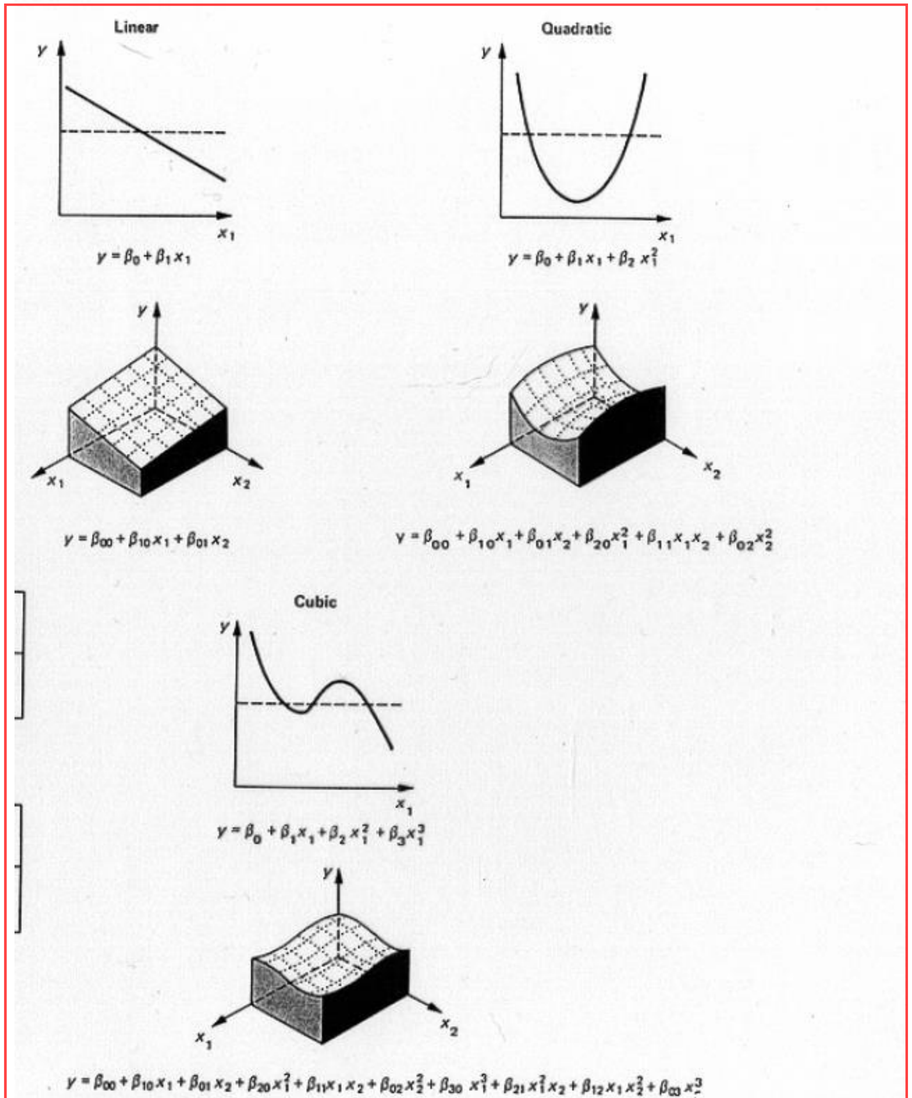
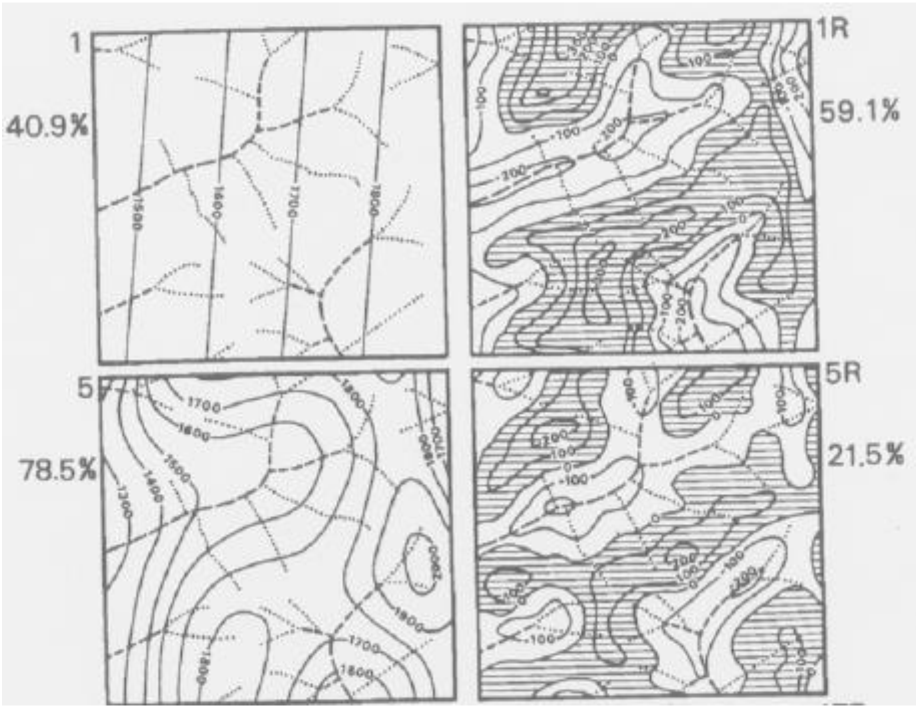


Figure 12.1 Relationship between two dimensional functions and their corresponding three dimensional surfaces. Source: Chorley and Haggett, 1965; Krumbein, 1956.

- The general equation of a trend-surface of order *p* is:
  $\hat{f}(\mathbf{s}_i) = \sum_{r+s \le p} b_{r,s} \cdot x_i^r \cdot y_i^s$,
  where *p* is an integer and the total number of parameters to be estimated is $(p+1)\cdot(p+2)/2$.
- The higher the polynomial order, the more complex surfaces can be modelled.
- The required order to model the general trend must be tested by a partial *F*-test. I.e., assuming that the coefficients of higher order terms are zero:
  - Comparing a 1^{st} order model against a 2^{nd} order model tests the null hypothesis:
    $$H_0: \beta_{20} = \beta_{11} = \beta_{02} = 0$$
  - Comparing a 2^{nd} order model against a 3^{rd} order model tests the null hypothesis:
    $$H_0: \beta_{30} = \beta_{21} = \beta_{12} = \beta_{03} = 0$$
- The coordinates at higher power will be highly correlated with the coordinates a lower power. See the script **`PolynomCorrelation.r`**.
- Furthermore, involving higher order powers of the coordinates may shift the numerical representation scale leading to a loss of numerical precision.
  This can be controlled to some degree by ***rescaling*** the coordinate system.
- We can (and should) compute the ***standard error*** $\sqrt{Var(\hat{y}_j)}$ of the predicted surface at any location $\mathbf{s}_j$:
  $$Var(\hat{y}_j) = \hat{\sigma}^2 \cdot \left((1, x_j, y_j, \ldots) \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot (1, x_j, y_j, \ldots)^T \right)$$
  This allows us to evaluate the prediction uncertainty.
  $\Rightarrow$ In general, towards the edge of the study area the uncertainty will increase because fewer points are located at the edge relative to the center of the study area.

Example: Erosional Terrain in Emporium Quadrangle, Pennsylvania



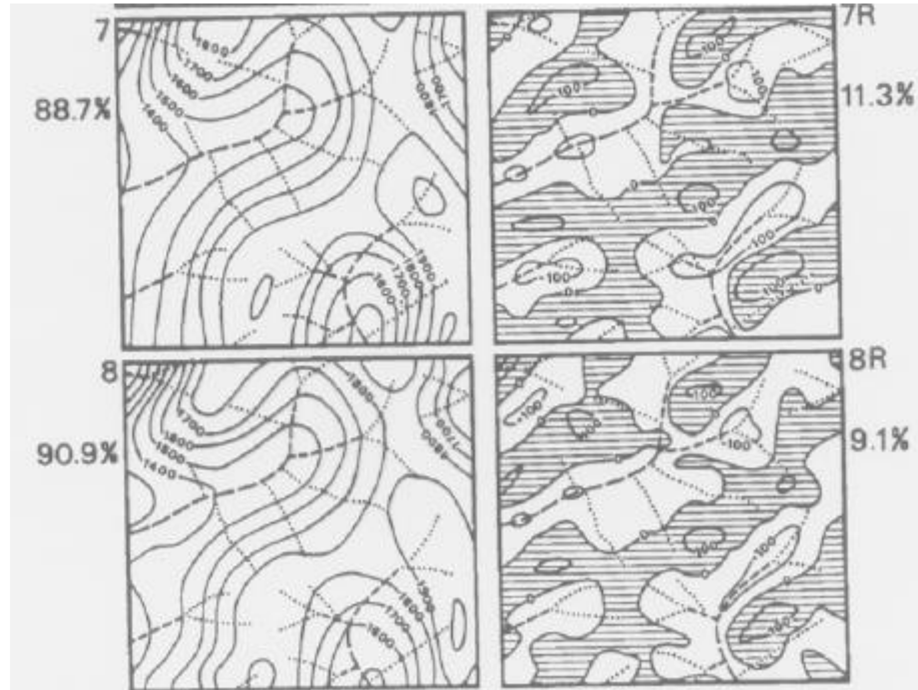
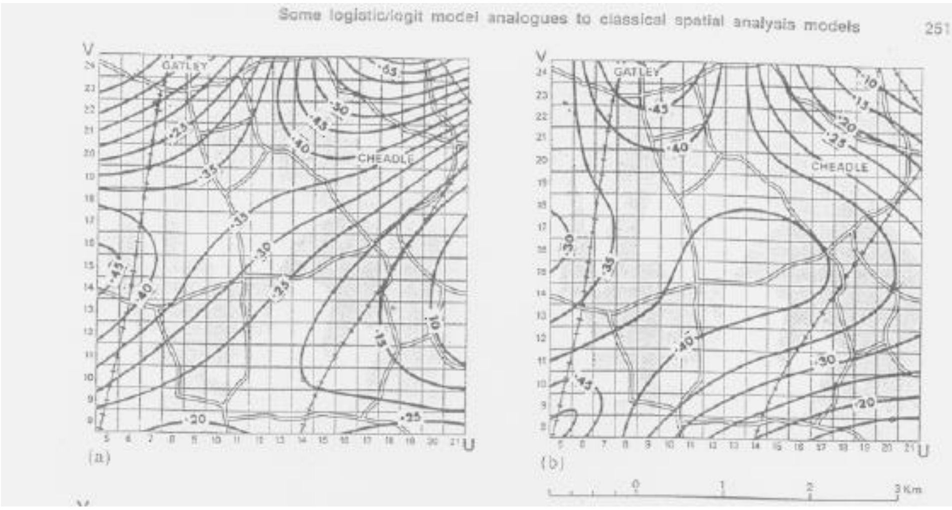
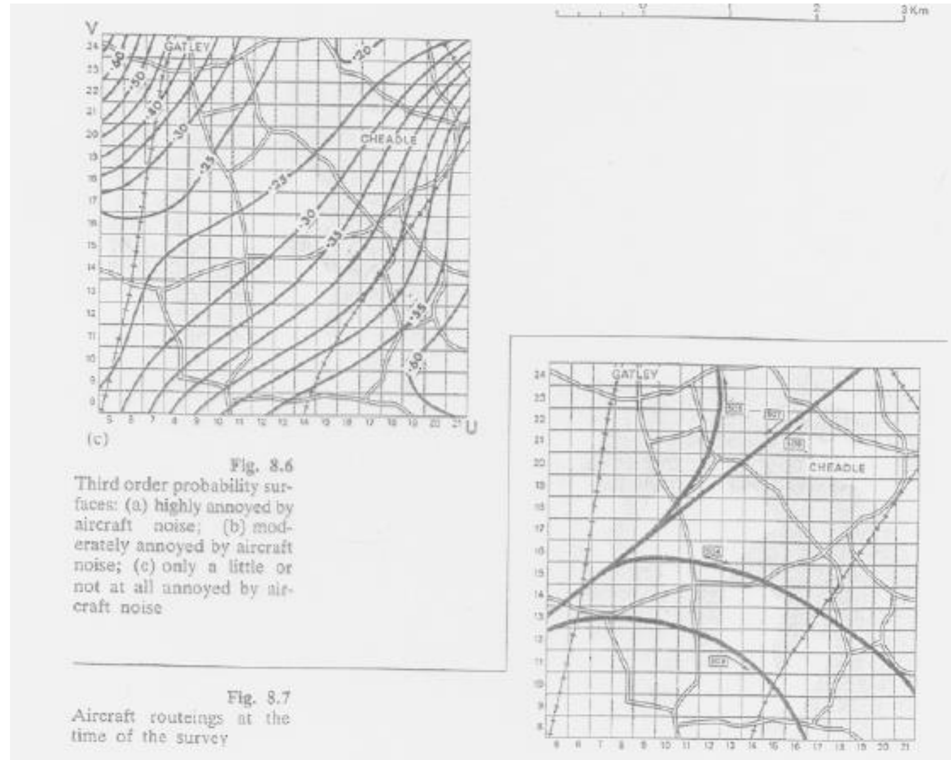


Fig. 6.9. Polynomial trend surfaces (*left*) and the associated residuals (*right*) for first, fifth, seventh, and eighth orders for the test region shown in Fig. 6.8. (Bassett and Chorley 1971)

Logistic regression trend-surface

- Logistic regression trend-surface can be used to model probability surfaces.
- The general form of a logit trend-surface is $\Pr(Y_i = 1) = 1/1 + \exp[-f(\mathbf{s}_i)]$
- The likelihood ratio test must be used to test for the order of the trend-surface.
- Wrigley's airport example:





- Note: Also harmonic surfaces can be modelled such as wave formation on a body of water. See Harbaugh & Preston, "Fourier Series Analysis in Geology", in Berry & Marble, *Spatial Analysis. A Reader in Statistical Geography*, **1968**, pp218-238

Notes of Caution

- The quality of trend-surface analysis dependents on the variability of the surface.
    - Smooth surfaces (such as climatological variables) can be modeled well.
    - Rugged surface (such as a mountainous terrain) are not appropriate for higher trend-surface modelling.

- Spatial heterogeneities can be captured by dummy variables.

- Global trend-surfaces cannot capture local variations
- The quality of trend-surface analysis dependents on the spacing of the training locations. Usually equally spaced training points give the best results.
- Higher order trend-surfaces become more complex and difficult to justify on theoretical grounds.
- Higher order trend-surfaces lead to multicollinearity among the polynomial transformed coordinates.
- An alternative approach to avoid multicollinearity is to use orthogonal polygons or two dimensional Fourier representations to model periodic surfaces.
  $\Rightarrow$ However, depending on the algorithm the value of the orthogonal polygon may not be available at the prediction locations.
- It is advisable to transform the coordinates to a fixed range (such as a z-transformation or $x_i \in [-1,1]$ and $y_i \in [-1,1]$) to avoid numerical instabilities and multicollinearity in estimating the regression parameters:
  - Centering around zero makes the design balanced (avoid multicollinearity to some degree)
  - Restricting the value range insures that higher order powers maintain the same numerical resolution.
- For large area interpolations the spherical distortion of the earth surface becomes relevant.
- Trend-surfaces should only be used for ***interpolation***, but should not be used for ***extrapolation*** outside the convex hull of the study region (edge problem). This is in particular relevant for higher order trend-surfaces.
- Trend-surface analysis assumes that the residuals are independent. Usually this assumption is violated for spatial data. The residuals, however, can be tested for spatial dependence.
- Feasible generalized least squares (FGLS) can accommodate trend-surface models with autocorrelated residuals.

Example: September Temperature in England, Wales and the Isle of Men

Table 5.17 The location and temperature of 23 weather stations in England, Wales, and the Isle of Man, together with their altitudes and the fitted values under two alternative models

Location	Coordinates X_{1i}	X_{2i}	Altitude (metres)	Mean September temperature (°C)	Fitted value for linear trend surface	Fitted value for augmented quadratic trend surface
Douglas	−0.162	0.179	85	15.2	16.29	15.60
Scarborough	0.109	0.190	52	18.1	17.91	18.48
Hull	0.110	0.130	2	19.3	18.21	19.07
Skegness	0.159	0.067	5	18.9	18.82	19.38
Lowestoft	0.258	−0.004	25	19.3	19.78	19.20
Clacton	0.216	−0.086	16	19.3	19.92	19.68
Buxton	0.009	0.076	307	16.0	17.85	16.75
Shawbury	−0.044	0.022	72	18.6	17.78	18.20
Malvern	−0.020	−0.056	62	19.3	18.32	18.90
Hurley	0.082	−0.117	43	19.9	19.25	19.78
Faversham	0.200	−0.140	48	20.4	20.09	19.69
Margate	0.240	−0.127	16	19.9	20.27	19.67
Hastings	0.180	−0.193	45	19.5	20.22	19.95
Squires Gate	−0.063	0.132	10	17.4	17.13	17.78
Colwyn Bay	−0.111	0.080	24	17.7	17.09	17.44
Loggerheads	−0.076	0.068	210	17.3	17.36	16.67
Lake Vyrnwy	−0.100	0.020	303	16.1	17.45	16.07
Aberporth	−0.175	−0.046	134	16.6	17.31	16.68
Swansea	−0.133	−0.106	10	18.3	17.86	18.48
Bude	−0.179	−0.196	15	18.2	18.02	18.49
Rosewarne	−0.235	−0.258	76	18.0	17.97	17.75
Durham	0.026	0.245	102	18.0	17.13	17.38
Poole	0.005	−0.202	5	19.9	19.18	20.20

Source: The Monthly Weather Report, September 1982: Meteorological Office.

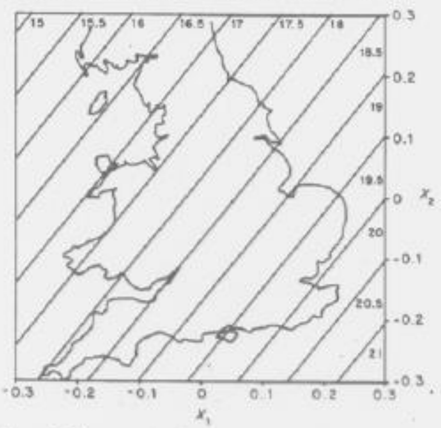


Figure 5.9 Linear trend surface fitted to September temperature data

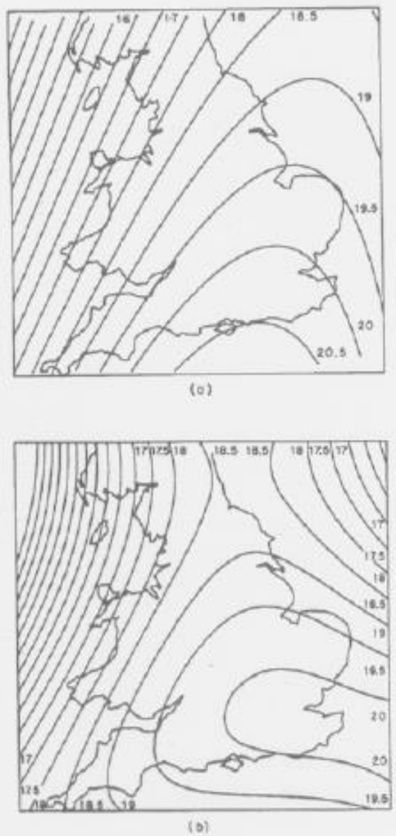


Figure 5.10(a) Quadratic (sea-level) trend surface fitted to September temperature data. (b) Cubic (sea-level) trend surface fitted to September temperature data

- See the R script **`TrendSurfaceModel.r`** for an analysis of these data.

Local Trend Surfaces

- Weighted least squares allows to fit a trend-surface to a local neighborhood.
- It also allows performing predictions.
- One way to define the weights is by the ***loess*** function: Let $q = \text{integer}(\alpha \cdot n)$ be a proportion of all data and δ the Euclidian distance of the q^{th} nearest point to the prediction location **s**. Then the weights are $w_i = \left[1 - \left(\frac{d(\mathbf{s},\mathbf{s}_i)^3}{\delta}\right)\right]_+^3$ where $[\ ]_+$ denotes $w_i = \begin{cases} w_i \text{ if } w_i > 0 \\ 0 \text{ if } w_i \leq 0 \end{cases}$.