

Spatial models and spatial modelling*

Roger Bivand

May 2002

*Talk prepared for CSISS spatial data analysis software tools meeting, Santa Barbara.

Outline

- Spatial models — a subset of models admitting spatial dependence among modelled objects/observations
- Relationships between observed data and hypothesised data generation processes
- How might we embed spatial models within the broader modelling paradigms in application domains?
- Can the choices we make about development modes of software for spatial modelling impact our ability to inter-operate with cognate modelling communities?

Spatial models

- Just position: point pattern, geometry, topology — input of prior knowledge about potential neighbours
- Univariate: relates to position through support
- Multivariate on same support (or maybe differing support that needs to be coarsed to same support)
- In univariate and multivariate cases: prediction and standard errors of prediction for same and (?) different support

Spatial dependence among observations

- Assembling and evaluating lists of neighbours sets — how real is the divide between geostatistical modelling and spatial econometric/lattice style modelling?
- Computing on lists of neighbours sets — computational geometry, topology
- From neighbours to weights, from sets to matrix representation
- How ought we to test between alternative specifications of spatial dependence? (just go with the one which gives us the results we want?)

Modelling with spatial position

- Data = Smooth + Rough
- Data = Global smooth + Local smooth + rough — geostatistics, maybe geographically weighted regression (anisotropy, non-stationarity)
- Data = $f(\text{Aspatial smooth}, \text{Spatial smooth})$ + rough — spatial econometrics, maybe multilevel models, Getis/Griffith-style filtering (non-stationarity again?)
- There should be added value if adding the spatial/local smooth relieves misspecification, scales back rough
- But we still have the classic issues of spatial data, potential non-stationarity, edge effects, etc

Relaxing assumptions about modelled variables

- Linear models with three-dimensional spatial data, maybe also 2D + time, even 3D + time
- Modelling variables relaxing assumptions about distributions — generalised linear models, robust models
- Non-linear modelling, local modelling in attribute and/or spatial domains, additive models
- Linear and non-linear mixed effects models, possibly including multilevel models

Modelling paradigms and applications domains

- Are the standard approaches to linking programs — integration, tight or loose coupling — still appropriate (XML, networked platforms, including GIS and databases)?
- How could we enhance the use of spatial modelling in the calibration of components of larger application suites (spatial modelling tools as context-sensitive scriptable plug-ins)?
- Perhaps rather than term the use of aspatial estimation techniques on spatial data potential misspecification, could we not rather share with others the opportunity spatial modelling offers to better understand and predict in the application context?

Computing on data, computing on models

- Using schema (following Chambers — classes) for models, model outcomes and/or model components may help to tease out the commonalities of spatial models (or their specificities)
- Using schema and methods dispatch then, by analogy with other models, lets us say `summary(model.obj)`, or maybe `plot(model.obj)`
- It becomes potentially more interesting if we say `anova(model.obj1, model.obj2)`, where `anova()` for our model class is an appropriate comparison of fit (say AIC or likelihood ratio)
- But how to say `predict(model.obj, new.data)` where `new.data` is a different map? What about `predict(model.obj, new.data, se=TRUE)`?

Development modes and opportunities for inter-operation

- The role of communities in building critical mass, both of users/developers, experience with analysis, and with ingestion of spatial data
- No necessary dichotomy between closed source/open source especially for scripted programming environments
- Acknowledge the importance of peer-review in reference implementations
- Recognise that “growing the cake” is important to the generation of demand for consulting and training in spatial data analysis
- Room for multiple delivery modalities, suited to the level of maturity and customisation of the tools