

Local Inferences via Nonparametric Model-Assisted Methods



F. Jay Breidt

Colorado State University

Jean D. Opsomer

Iowa State University



FUNDING SOURCE

- This presentation was developed under the STAR Research Assistance Agreement CR-829095 awarded by the U.S. Environmental Protection Agency (EPA) to Colorado State University. This presentation and the following have not been formally reviewed by EPA. The views expressed here are solely those of its authors and the STARMAP Program. EPA does not endorse any products or commercial services mentioned in this presentation.



Scales of Local Inference in Surveys

- Large area:
 - sample itself suffices for inference
 - no model needed
- Medium area:
 - use auxiliary information through a model
 - model helps inference but is not critical
- Small area:
 - sample size is small or zero
 - inference must be based on a model



Specific and Generic Inference

- Specific: one study variable, few population parameters
 - lots of modeling resources to specify, estimate, and diagnose a model
 - willingness to defend the model
- Generic: many study variables, many population parameters
 - no resources to model every variable
 - no single model is adequate/defensible



Generic Inferences in Aquatic Resources

- Generic inference is a common problem for federal, state, and tribal agencies
- Example: conduct a survey and prepare a report
 - analyze large numbers of chemical, biological, and physical variables
 - estimate means, quantiles, and distribution functions
 - break down both by political classifications and by various ecological classifications



Model-Assisted Survey Inference

- Scarce modeling resources for generic inference, so we don't trust models
- Can we use a model without depending on the model?
- Model-assisted inference:
 - efficiency gains if model is right
 - sensible inference even if model is wrong



Model-Assisted Estimators

- Form of model-assisted estimator:
 - (model-based prediction)+(design bias adjustment)
 - model incorporates auxiliary information
 - bias adjustment corrects for bad models
- Classical parametric model-assisted:
 - prediction from linear regression model
- Our idea: nonparametric model-assisted
 - prediction from kernel regression or other “smoother”



Why Nonparametric?

- More flexible model specification
 - smooth mean function, positive variance function
- Approximately correct more often
 - more opportunities for efficiency gains from auxiliary information
 - often, not a large efficiency loss if parametric specification is correct



Goals of Our Research

- Incorporate auxiliary information for small or medium-scale local inference
- Focus on generic inference
- Use flexible nonparametric models to reduce misspecification bias
- Make the methods operationally feasible for state and tribal agencies



Future Directions

- Adapt the methodologies to new design contexts:
 - multiple phases, multiple stages
- Investigate alternative smoothing methodologies:
 - semiparametric and additive models
- Prove theorems:
 - smoothing parameter selection



Outline of Presentations

- Siobhan Everson-Stewart
 - “Nonparametric Survey Regression Estimation in Two-Stage Spatial Sampling”
- Alicia Johnson
 - “Estimating Distribution Functions from Survey Data Using Nonparametric Regression”
- Mark Delorey
 - “Semiparametric Mixed Models in Small Area Estimation”