From neuroscience to national science

Three problems across the union of geometry and policy

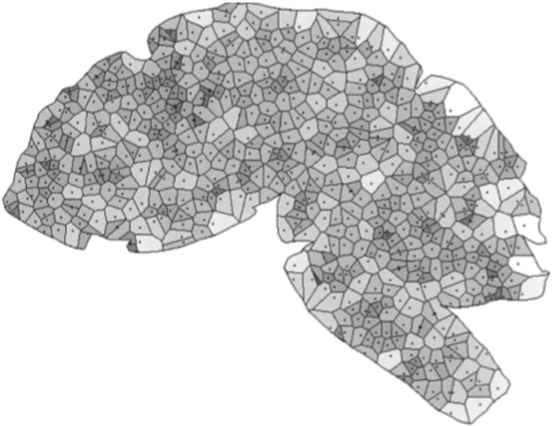
Christopher D. Barr, Ph.D.

Harvard Biostatistics

Outline

- Voronoi estimator
  - Investigate and improve the Voronoi estimator
    - Derivations of sampling distribution and bias
    - Foundational result in stochastic geometry
- Optimal allocation of resources
  - Find the optimal placement of facilities
    - Versatile driver for solving a range of variants
    - Connect image analysis and operations research
- Climate change
  - Improve climate change legislation and air pollution mitigation
    - Draw a connection and strengthen both cases
    - Prime projects estimating gains from ancillary mitigation

Section 1: The Voronoi estimator

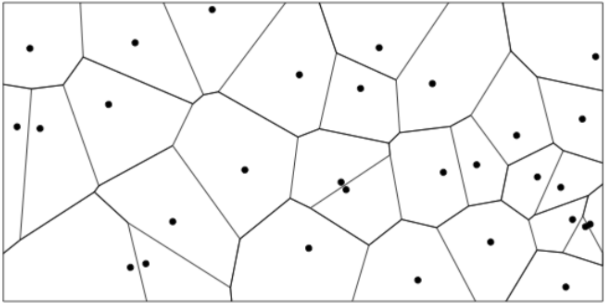


- Joint with Rick Schoenberg -

Definition of a Voronoi diagram

Cell C_i contains all locations closer to point p_i than any other point

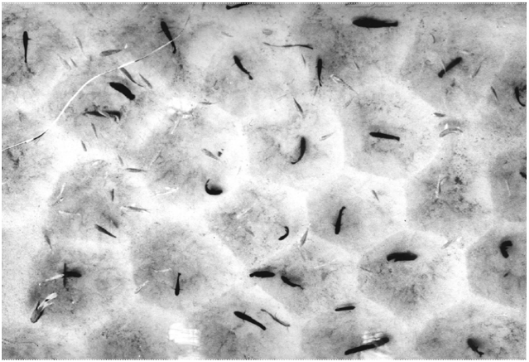
$$C_i = \{y \in \mathcal{S} : ||y - p_i|| \leq ||y - p_j||, \forall p_j \in N\}$$



Practical examples

Voronoi diagrams have a natural interpretation as territories

- Breeding regions of animals
- Cells growing in a dish at a constant rate



Future Voronoi tessellation

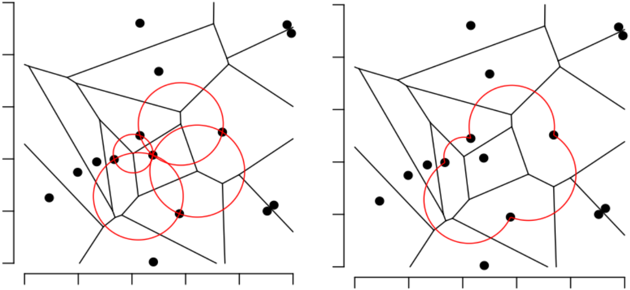
Photo courtesy of Travis Gerke, 12-25-2009

Text message: "Future Voronoi tessellation"

Fundamental domain

Points outside the fundamental domain have no influence

$$FD(C_i) = \bigcup_{y \in C_i} B(y; ||y - p_i||)$$



Motivation and goals

Voronoi estimator proposed

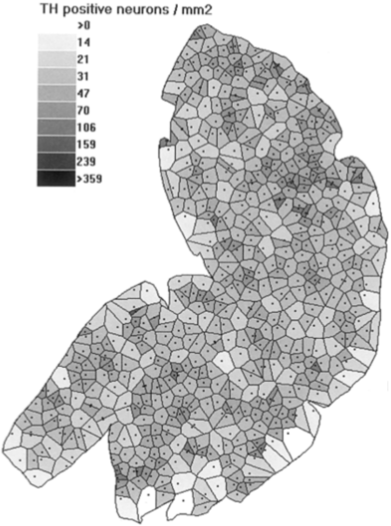
- Neuroscience method
- Limited precedent

Theoretical goals

- Bias
- Sampling distribution

Practical goals

- Improvements
- Applicability



A framework in stochastic geometry

The classic pursuit

- Draw a realization from a point process
- Tessellate the point pattern
- Record a property associated with a random cell

Recent efforts

- Moment type results in closed form
- Distributional results largely by simulation

Two types of cells

- Typical cell, randomly selected cell
- Voronoi estimator cell, cell covering a specific location

Results for the typical homogeneous Poisson Voronoi cell

Closed form

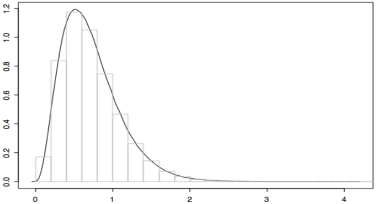
Expected cell area is λ^{-1}

Expected cell perimeter is $4\lambda^{-1/2}$

Expected number of sides is 6

Simulation

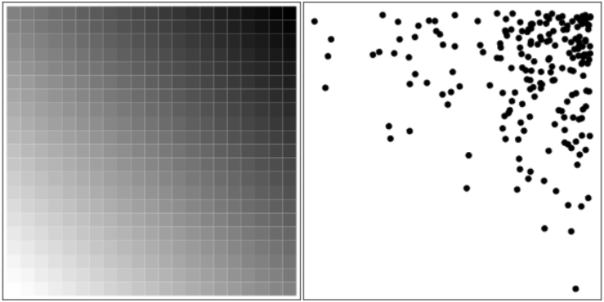
Distribution of cell area is gamma, $\mathcal{A} \sim G(x; \alpha, \beta)$



Inhomogeneous Poisson point process

Regulate by an intensity function $\lambda(y)$

$$Pr\{N(A) = i\} = \frac{[\int_A \lambda(y) d\mu]^i}{i!} exp\left\{- \int_A \lambda(y) d\mu\right\}$$



Assumptions

Area of the typical inhomogeneous Poisson Voronoi cell is gamma

Assumption 1: $\mathcal{A}_{t(y)} \sim G(x; \alpha_y, \beta_y)$

Intensity is bounded globally

Assumption 2: $M_1^{(k)} < \lambda^{(k)}(r, \theta) < M_2^{(k)}, \forall (r, \theta) \in \mathcal{S}$

Intensity is locally homogeneous

Assumption 3: $\lambda^{(k)}(r, \theta) = a_{(k)}$, if $r < b_{(k)}$

Intensity is locally homogeneous

Assumption 4: $\frac{\exp\{-\pi M_1^{(k)} b_{(k)}^2 / 4\}}{M_1^{(k)}}$ converges to 0 as $k \to \infty$

Theoretical results

Expected area of the typical cell is the reciprocal of the intensity

Theorem 1: $E(\mathcal{A}_{t(y)}) \rightarrow 1/\lambda_y$

Sampling distribution of the Voronoi estimator is inverse gamma

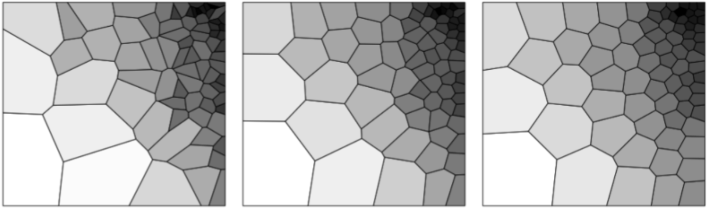
Theorem 2: $\hat{\lambda}_y \sim IG(x; \alpha_y + 1, 1/\beta_y)$

Voronoi estimator is approximately unbiased

Corollary 1: $E(\hat{\lambda}_y) \rightarrow \lambda_y$

Centroidal Voronoi estimator

- Points moved to centroid at each iteration
  - Will smooth to homogeneous if allowed



Kernel estimators

Kernel intensity estimation

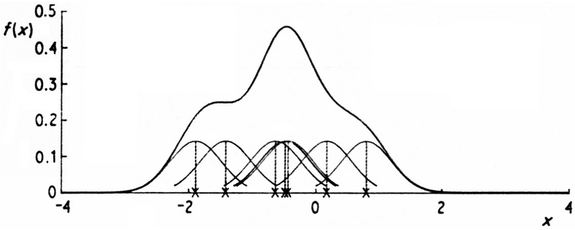
$$\lambda^k(y;\gamma,S) = \left\{\sum_{i=1}^{n} k(||y - y_i||;\gamma)\right\}/p(y;\gamma,S).$$

Gaussian kernel is common

Fixed and adaptive bandwidth techniques

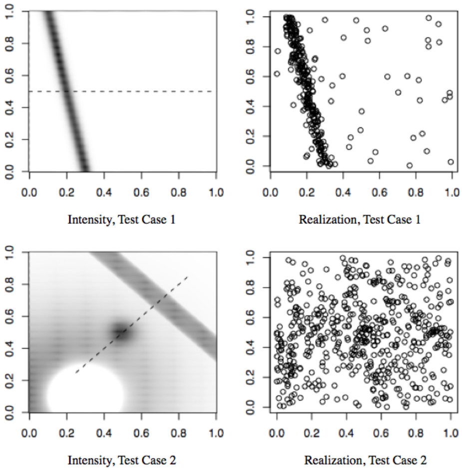
Drawbacks

Bias, parameter selection



Two test cases

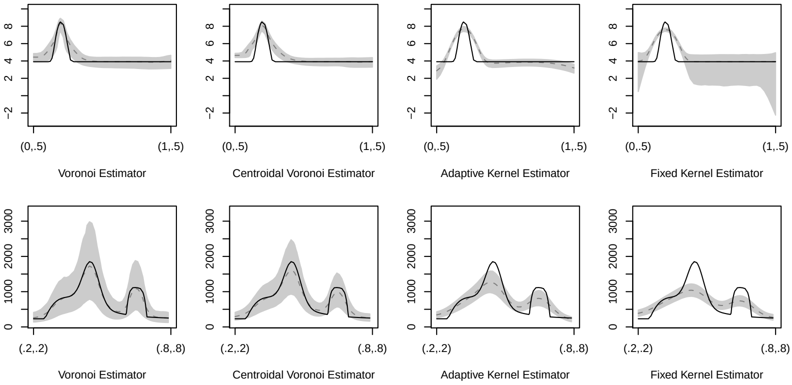
- Test Case 1
  - Extreme variation
  - Example to follow
- Test Case 2
  - Heikkinen and Arjas
  - Bayesian Voronoi



Performance

Comparison along one hundred locations in each test case

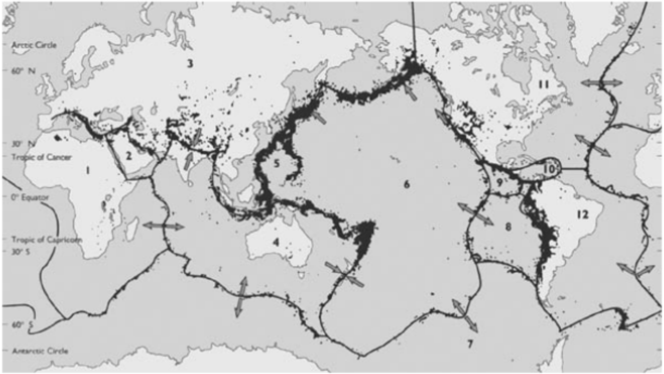
Centroidal Voronoi estimator trades bias for reduced variance



Introducing earthquakes

Global earthquakes with tectonic plates

High intensity along many plate boundaries



Earthquake intensity

Original data set
- Magnitude > 3
- Since January 1977

Voronoi estimator
- Reasonable in the East
- Responsive in the West

Kernel estimator
- Zero in the East
- Too smooth in West

Large magnitude earthquakes

Voronoi estimator

Fixed bandwidth kernel estimator

Summary

Challenge

Investigate and improve the Voronoi estimator

Contributions

Derivations of sampling distribution and bias

Foundational result in stochastic geometry

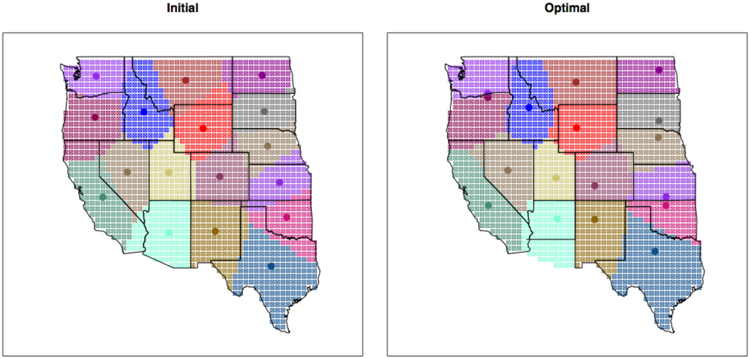
Future work

Iterations for centroidal estimator, neuroscience collaboration

Reference

Barr CD, Schoenberg FP. On the Voronoi estimator for intensity of an inhomogeneous planar Poisson process. *Biometrika*, tentatively accepted

Section 2: Optimal allocation of resources



- Joint with Travis Gerke -

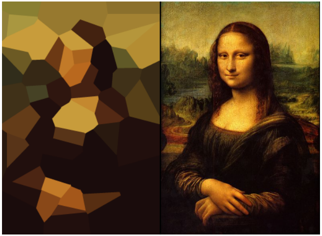
Possible approaches

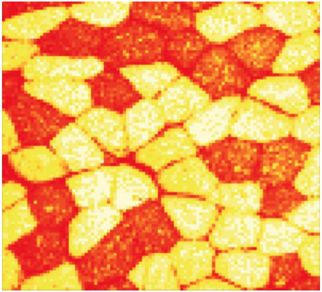
Operations research

- Placement of schools
- Gradient descent

Bayesian image analysis

- Model locations and colors
- Bayesian framework
- MCMC to search space





Introducing the model

Locations and colors as the parameter space to be searched

$$\Theta = \{(u_i, v_i, c_i), i = 1, 2, ..., n\}$$

Starting with an uninformative prior

$$\Pi(\Theta) = \Pi_{i=1}^{n} U(0,1)^2 \times p_c^{I(c_i=\eta_1)}(1-p_c)^{I(c_i=\eta_{-1})}$$

Suppose Gaussian error

$$P(z|\Theta) \propto exp\left\{-\frac{1}{2\sigma^2}\Sigma_{j,k=1}^{N}(z_{jk} - \mu_{jk}(\Theta))^2\right\}$$

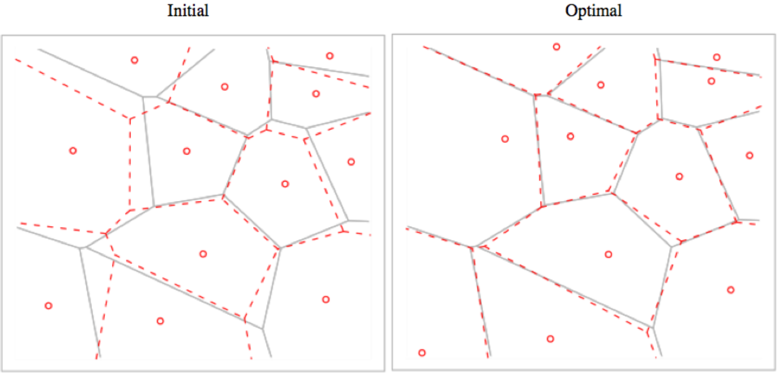
Posterior to be explored

$$P(\Theta|z) \propto \Pi(\Theta) \times P(z|\Theta)$$

Recovering a tessellation

Forming optimal Voronoi covers

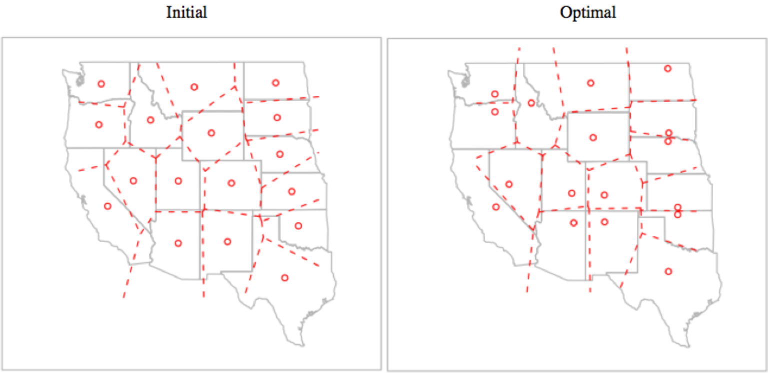
Regions that are Voronoi cells can be matched exactly



Federal emergency response centers

Find the best location for one center per state

Notice Washington and Oregon, Kansas and Oklahoma



Flexible framework

A solution more locally optimal

Inhibitory prior on the location of points

Variable number of facilities

Reversible jump MCMC

Initial Optimal

Summary

Challenge

Find the optimal placement of facilities

Contributions

Versatile driver for solving a range of variants

Connect image analysis and operations research

Future work

Theoretical underpinnings, broader applicability of driver

Reference

Barr CD, Gerke TA. Allocating point-type resources
to form optimal Voronoi covers. *In progress*.

Section 3: Climate change

- Joint with Francesca Dominici -

The ACES Act

- American Clean Energy and Security Act (ACES) of 2009
  - Regulates carbon dioxide, methane, nitrous oxide and others
  - Capped at 2005 annual levels and reduced by 2% per year
  - Buy and sell pollution permits
  - Tariffs on countries that fail to pass similar legislation
- Passed in the House
  - June 26, 2009 by a vote of 219-212
- Debate
  - Economic costs and implied tax across the country
  - Mitigation of global warming, associated public health issues

Shared sources

Co-pollutant: NO_x

NO_x	**21,450**	**21,070**	**19,004**	**15,612**	**14,701**	**14,250**
Mobile Fossil Fuel Combustion	10,920	10,622	10,310	8,757	8,271	7,831
Stationary Fossil Fuel Combustion	9,689	9,619	7,802	5,857	5,445	5,445
Industrial Processes	591	607	626	534	527	520
Oil and Gas Activities	139	100	111	321	316	314
Incineration of Waste	82	88	114	98	98	97
Agricultural Burning	28	29	35	39	38	37
Solvent Use	1	3	3	5	5	5
Waste	0	1	2	2	2	2

Greenhouse Gas: CO_2

Gas/Source	1990	1995	2000	2005	2006	2007
CO_2	**5,076.7**	**5,407.9**	**5,955.2**	**6,090.8**	**6,014.9**	**6,103.4**
Fossil Fuel Combustion	4,708.9	5,013.9	5,561.5	5,723.5	5,635.4	5,735.8
Electricity Generation	1,809.7	1,938.9	2,283.2	2,381.0	2,327.3	2,397.2
Transportation	1,484.5	1,598.7	1,800.3	1,881.5	1,880.9	1,887.4
Industrial	834.2	862.6	844.6	828.0	844.5	845.4
Residential	337.7	354.4	370.4	358.0	321.9	340.6
Commercial	214.5	224.4	226.9	221.8	206.0	214.4

Quantifying the ancillary benefits of cap and trade

- Reduction in air pollutants associate with cap and trade
  - Scenarios with different sectors
  - National emissions inventory
- Connecting air pollution and human health
  - Top ten public health problems by World Health Organization
  - 3000 hospital admissions for every 10 $\mu g/m^3$
  - Half a year of life lost for every 10 $\mu g/m^3$

Summary

Challenge

Improve climate change legislation and air pollution mitigation

Contributions

Draw a connection and strengthen both cases

Prime projects estimating gains from ancillary mitigation

Future work

Quantify the benefits of ancillary air pollution mitigation

Reference

Barr CD, Dominici F. Cap and trade legislation for greenhouse gas emissions: additional public health benefits from air pollution mitigation. *JAMA* (2010).

Conclusion

- Voronoi estimator
  - Investigate and improve the Voronoi estimator
    - Derivations of sampling distribution and bias
    - Foundational result in stochastic geometry
- Optimal allocation of resources
  - Find the optimal placement of facilities
    - Versatile driver for solving a range of variants
    - Connect image analysis and operations research
- Climate change
  - Improve climate change legislation and air pollution mitigation
    - Draw a connection and strengthen both cases
    - Prime projects estimating gains from ancillary mitigation

From neuroscience to national science

Three problems across the union of geometry and policy

Christopher D. Barr, Ph.D.

Harvard Biostatistics