

Development of an Automated Pesticide Exposure Analyst for the California's Central Valley

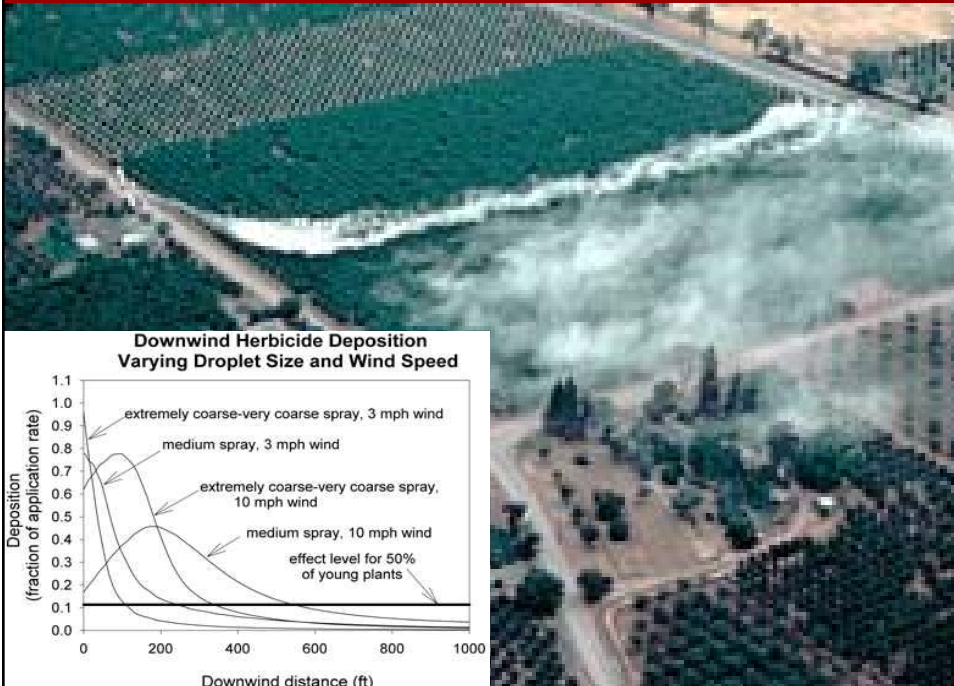
Daniel W. Goldberg
GIS Research Laboratory
Department of Computer Science
University of Southern California
<http://uscgislab.net>



Pesticides and prostate cancer

- Studies rely on
 - self-reported occupational exposure
 - self-reported home exposure
 - assumed exposure for an occupation
 -all resulting in bias towards the null
 - assess only current exposure
- omission of non-occupational ‘ambient’ environmental exposures (e.g. residences downwind of application sites)

Why is ambient exposure important?



California Environmental Protection Agency
 **Air Resources Board**

**Ambient Air Monitoring
for Methyl Bromide and 1,3-Dichloropropene
in Kern County – Summer 2001**

Concentrations of methyl bromide were measured as high as 98.3 micrograms per cubic meter of sampled air ($\mu\text{g}/\text{m}^3$) at the University of California's Cotton Research Station (CRS) near the town of Shafter. Methyl bromide was used north and northwest of the CRS site associated with growing roses. The prevailing wind in the CRS area is from the northwest. No methyl bromide was used at the CRS during 2001. The highest average concentration for the 8-week canister monitoring period was $11 \mu\text{g}/\text{m}^3$, also measured at the CRS site. Of the 198 ambient canister samples, 171 contained concentrations of methyl bromide above the estimated quantitation limit (reporting limit) of $0.036 \mu\text{g}/\text{m}^3$, and 27 samples were invalidated due to a sampling problem.

Pesticides/herbicides in California

- California accounts for approximately 13% (\$957 million) of all agricultural chemical expenditures in the U.S., including pesticides
- In 2000, approximately 172 million pounds of pesticide active ingredients were applied for production agriculture in California

Epidemiological Challenges

- Nobody knows how much pesticide they are ‘ambiently’ exposed to
 - “please estimate how much oxychlordan you have inhaled in the last 2 years”
- Biomarkers are hard to obtain and costly
- Current exposures are probably not relevant to disease onset

Aim

- Develop a system for deriving an estimate of **lifetime** history of **ambient exposure** to pesticides for California’s Central Valley
- **A LOT** of related work, e.g.,
 - California
 - Bell et al. 2001 – Fetal Death
 - Rull et al. 2001, 2003 – Parkinson’s Disease
 - Reynolds et al. 2005 – Breast Cancer
 - Nuckols et al. 2007 – Exposure Estimation
 - Massachusetts
 - Brody et al. 2002, 2004 – Breast Cancer
 - Nebraska
 - Ward et al. 2000 – Non-Hodgkin Lymphoma

Technical Goal

- Can we build an exposure estimation system which is:
 - Flexible
 - Types of exposure classification areas
 - Spatial buffers, census blocks
 - Types/Accuracies of exposure level data sources
 - PLSS, LU, PUR
 - Types of exposure calculations
 - Constant densities, tiered densities
 - Scalable
 - Any number of subjects
 - Any number of years
 - Any number of chemicals
 - Simple to re-run as better data/calculations become available

Technical Challenges

- Data sources are rarely available tailored directly to a study's needs
 - Varying characteristics amongst data sources
 - Temporal
 - Resolution
 - Accuracy
 - Spatial
 - Resolution
 - Accuracy
- Large populations result in high computational complexity

Test Area

- California's Central Valley
 - Fresno, Tulare, Kern Counties

- Included in 4 studies with different:
 - Methodologies
 - Data Sources
 - Exposure Calculations

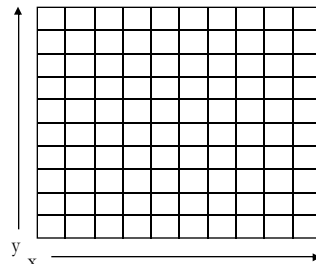
- **Bell et al. 2001, Rull et al. 2001, 2003, Reynolds et al. 2005, Nuckols et al. 2007**



Data Sources

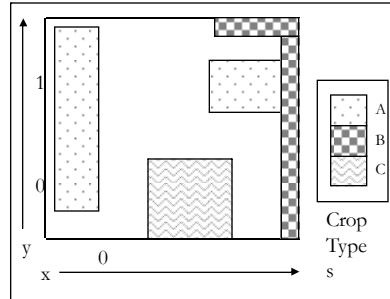
- Pesticide Use Reports (PUR)
 - In 1972, California mandated by law the filing of for commercial applications of restricted-use pesticides (i.e., agents with harmful environmental or toxicological effects)
 - Extended to cover all pesticides in 1990
 - Data available digitally from 1972

–*Problem:* pesticide applications are reported according to the *Public Land Survey System (PLSS)*, sections with an area of approximately **1 mi² (640 acres or 259 ha)**

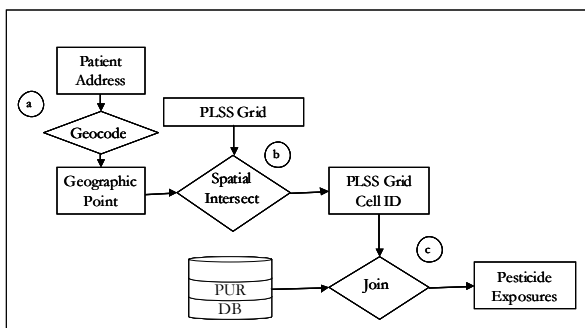


Data Sources

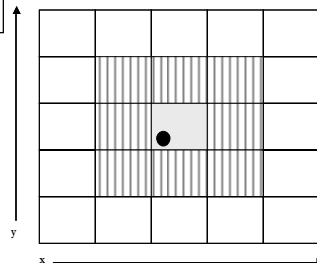
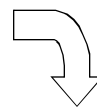
- Land-Use (LU) Survey Data
 - Available from the California Department of Water Resources
 - Performs countywide, large-scale surveys (1:24,000, or 1 in. = 2,000 ft) of land use and crop cover every 7–10 years
 - Timing varies by county (staggered)
 - Data back to 1952, digitized manually back to 1969



Bell et al. 2001

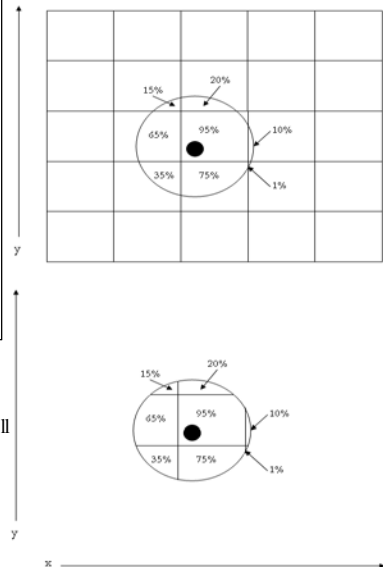
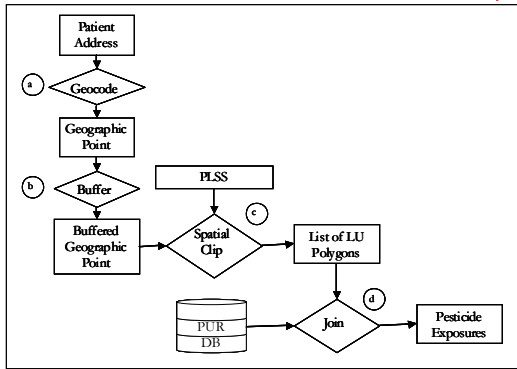


- Classify as:
 - Exposed/Unexposed
 - Broad/Narrow



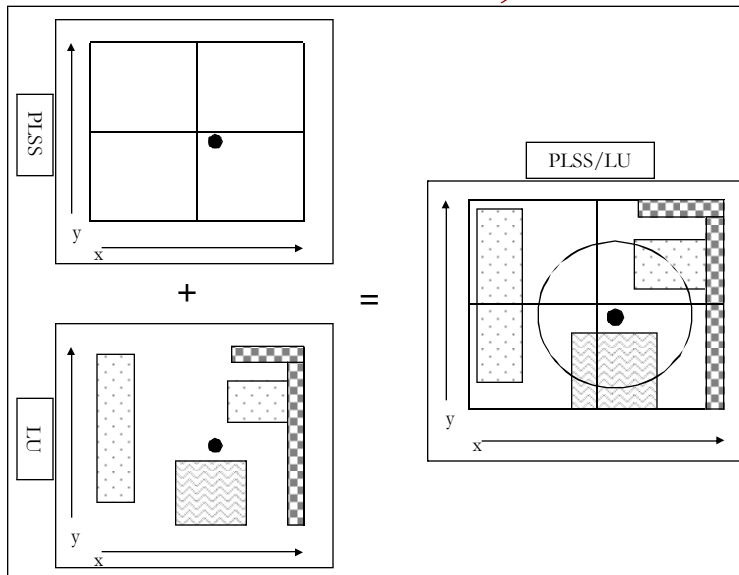
- Application Density = PUR Density
- Exposure = PUR Pounds * # PLSS Cells

Bell et al. 2001, Extended

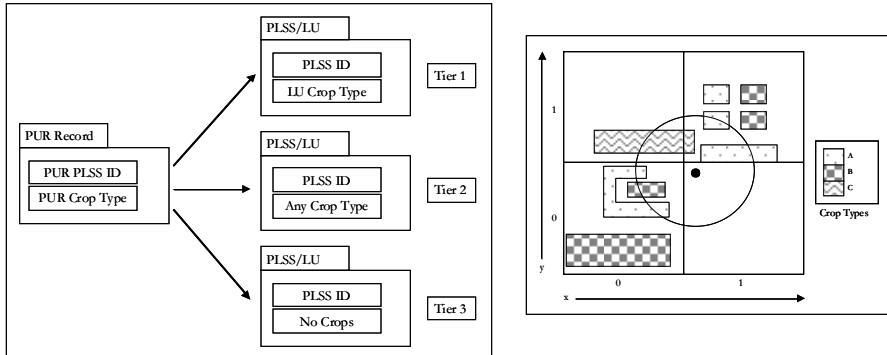


- Application Density Per Cell = PUR Density
- Exposure Per Cell = Application Density Per Cell / Area of Cell
- Exposure = Sum(Exposure Per Cell)

Rull et al. 2001, 2003

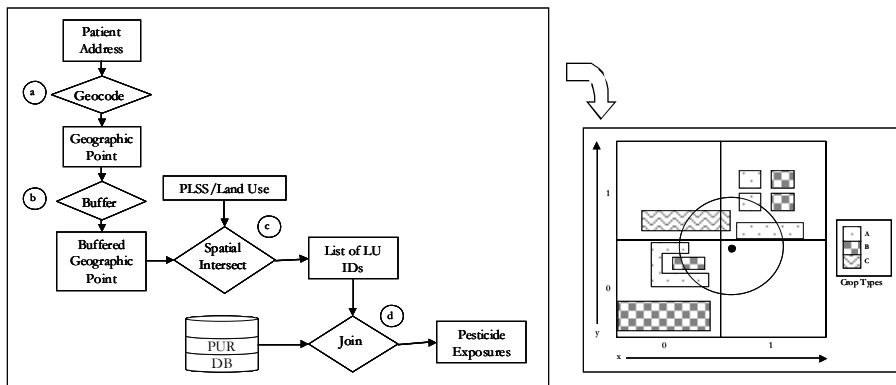


Rull et al. 2001, 2003



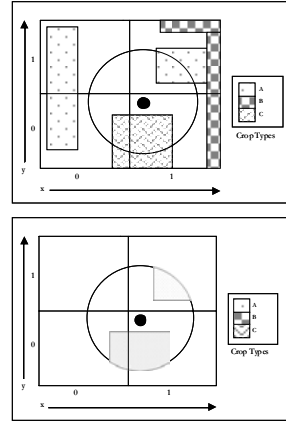
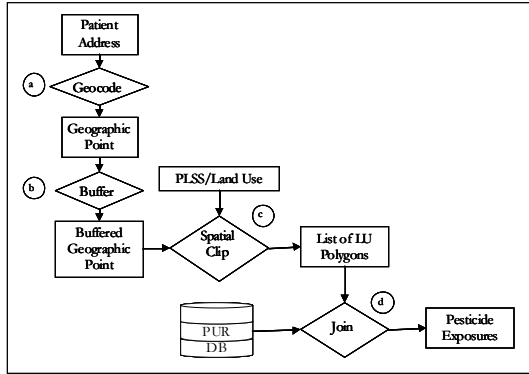
Row	Chemical	Cell	Crop	Pounds	Acres	Tier	Fields	Intersects
1	a	[1,0]	A	25	300	3	0	0
2	c	[0,0]	D	15	30	2	3	3
3	b	[0,0]	B	10	100	1	2	1
4	a	[0,0]	A	20	60	1	1	2

Rull et al. 2001, 2003



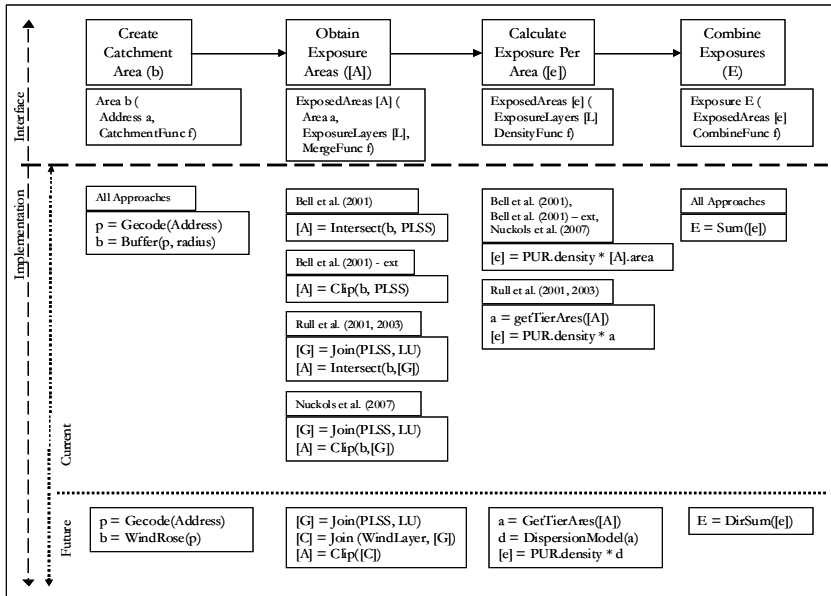
- Application Density Per Cell = PUR Density
- Exposure Per Cell = Tiered Exposure
- Exposure = Sum(Exposure Per Cell)

Nuckols et al. 2007



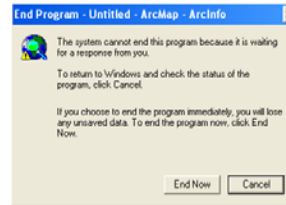
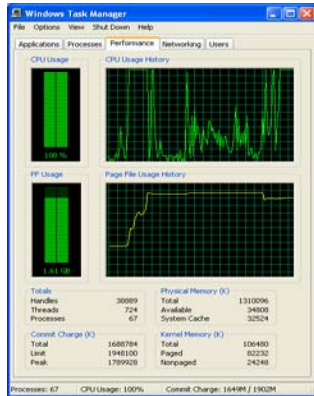
- Application Density Per Cell = PUR Density
- Exposure Per Cell = Application Density Per Cell / LU polygon
- Exposure = Sum(Exposure Per Cell)

Generalization



Scalability

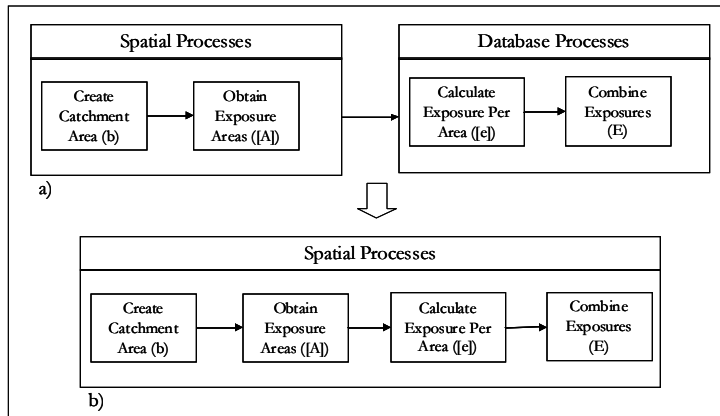
- As all variables increase, so does the complexity:
 - # subjects, #years, #chemicals, etc.
- Need to be able to scale reliably



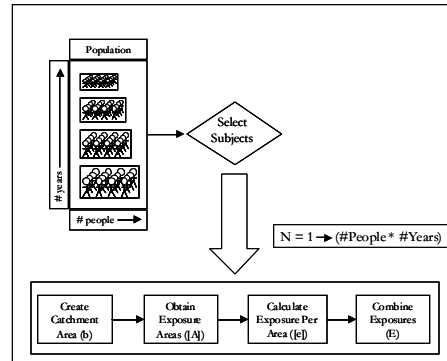
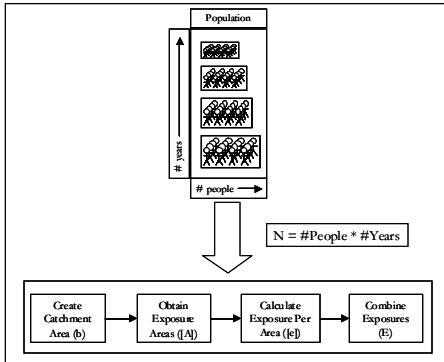
Scalability – Processing Location

Bell et al. (2001): N=684
 Rull et al. (2001, 2003): N=200
 Nuckols et al. (2007): N=577

“This approach was too computationally intensive for a simulation exercise involving more than 77,000 parcels“ (Ritz, pc 2007)

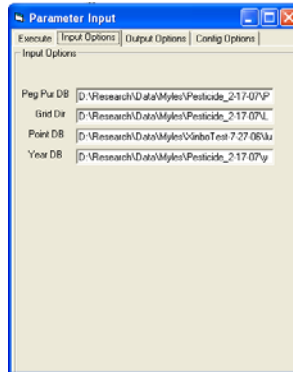
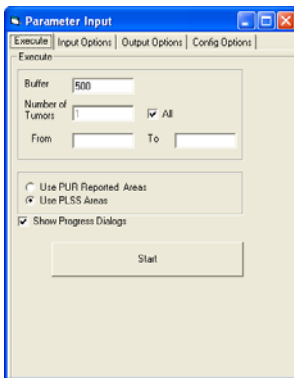


Scalability – Per-Subject Processing

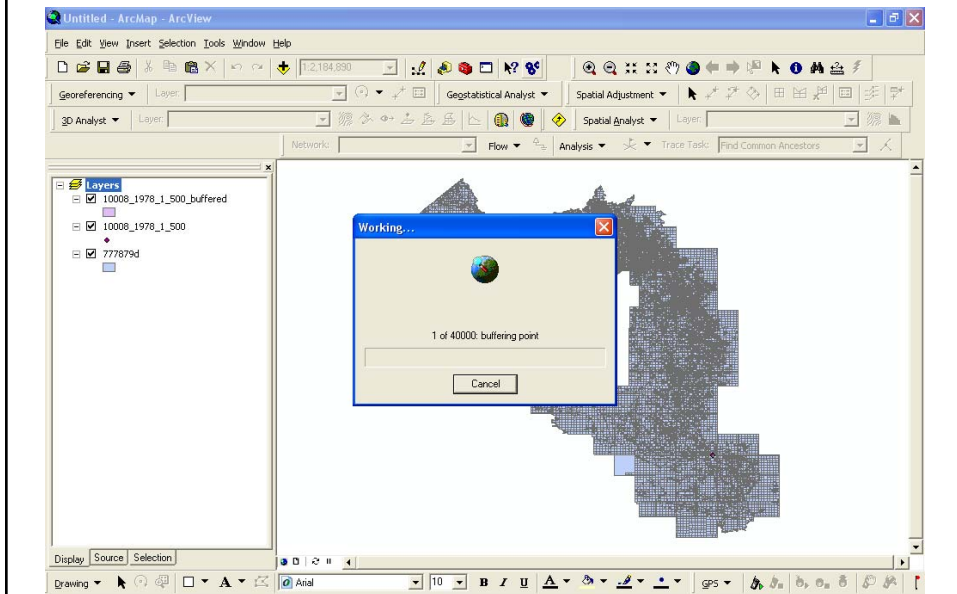


Bell et al. (2001): $N=684$
 Rull et al. (2001, 2003): $N=200$
 Nuckols et al. (2007): $N=577$

Prototype



Prototype



Summary

- Many ways to estimate exposure depending on
 - Data available
 - Calculations used
- By generalizing the process, we can build tools which are:
 - Automated
 - Scalable to the sizes of populations needed
 - Extensible as new data/calculations are developed