Geographical Clusters of RSEI Hazard Scores in the USA

Counties → Census Tracts → Block Groups → Blocks

Raid W. Amin, University of West Florida, ramin@uwf.edu
Toni Charlot, University of West Florida, tc30@students.uwf.edu

2016 National Training Conference on the Toxics Release Inventory and Environmental Conditions in Communities, 19-21 October 2016, Washington DC, USA

Introduction
About SaTScan
The Zoom-in Approach

RSEI
About the Data
About the study
Carcinogenic pollution Rough Trend 2000-2011
Zoom-in
The Northern Illinois Cluster

Conclusion
Final Remark

Spatial analysis using SaTScan[1]

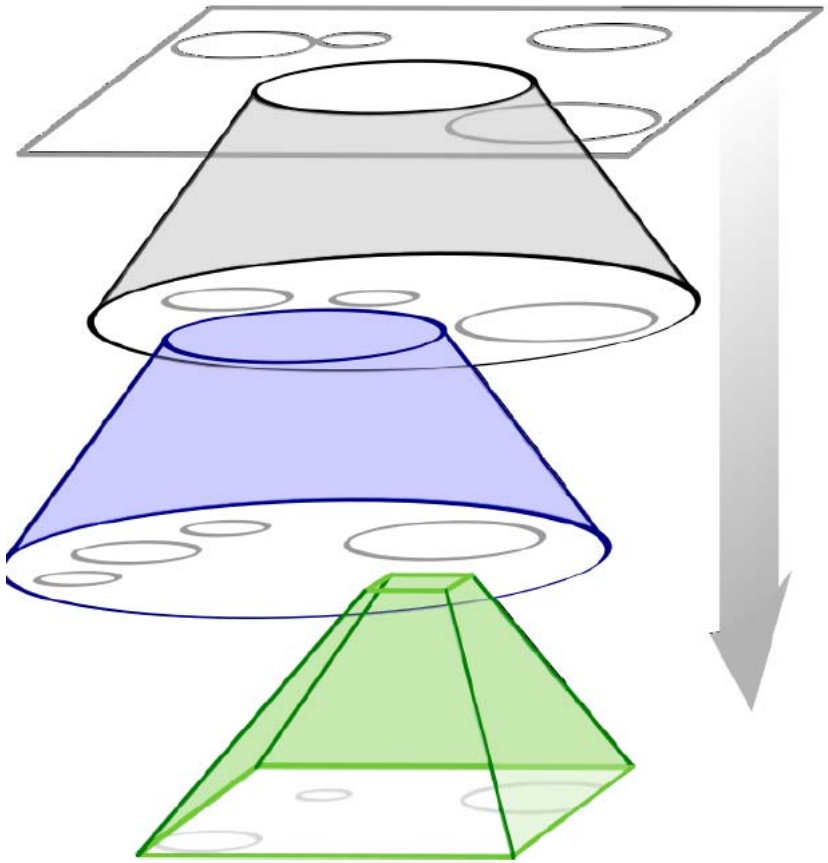
- Free for use, can be downloaded at www.satscan.org
- We used the latest version, v9.4.1, March 23rd, 2015.
- We utilized the Normal Model for continuous data.
- Maximizes the likelihood function over circles of varying sizes. We used a 5% window size.
- A circle with the highest log likelihood ratio (LLR) is least likely to have occurred by chance.
- Cluster p-value obtained using 9,999 Monte Carlo replications.
- In this study, p-value < 0.05 is used to determine significant clusters.

1 Developed by Martin Kulldorff

The Zoom-in Approach

- Aggregate large geo-datasets onto larger geographical regions.
- Perform cluster analysis.
- Select clusters of interest.[1]
- Get data for smaller geographical areas within selected clusters.[2]
- Go back to step 2 until the smallest geographical area is reached.

1. Clusters with the highest log likelihood ratio, highest/lowest mean inside or those in a particular location of interest.
2. Normalize the data for smaller geographical areas on the whole set prior to extracting the smaller dataset of interest in order to maintain an unambiguous reference to the larger context.



Risk Screening Environmental Indicators(RSEI)[4][3]

- A computer-based screening tool developed by the EPA to analyze risk factors and put the Toxics Release Inventory (TRI) database into a chronic health context.
- The EPA uses km^2 grid cells ($101km \times 101km$) around each facility and estimates exposure for each chemical released[2].
- In other words, each facility generates 10,200 entries for each chemical released, for each media such as **landfills, air, streams or wells**. The TRI dataset contains releases for over 600 chemicals from over 70,000 facilities or sites.
- The RSEI approach augments estimates of pounds released with toxicity and exposure considerations[3].

Illustration of km^2 Grid Cells

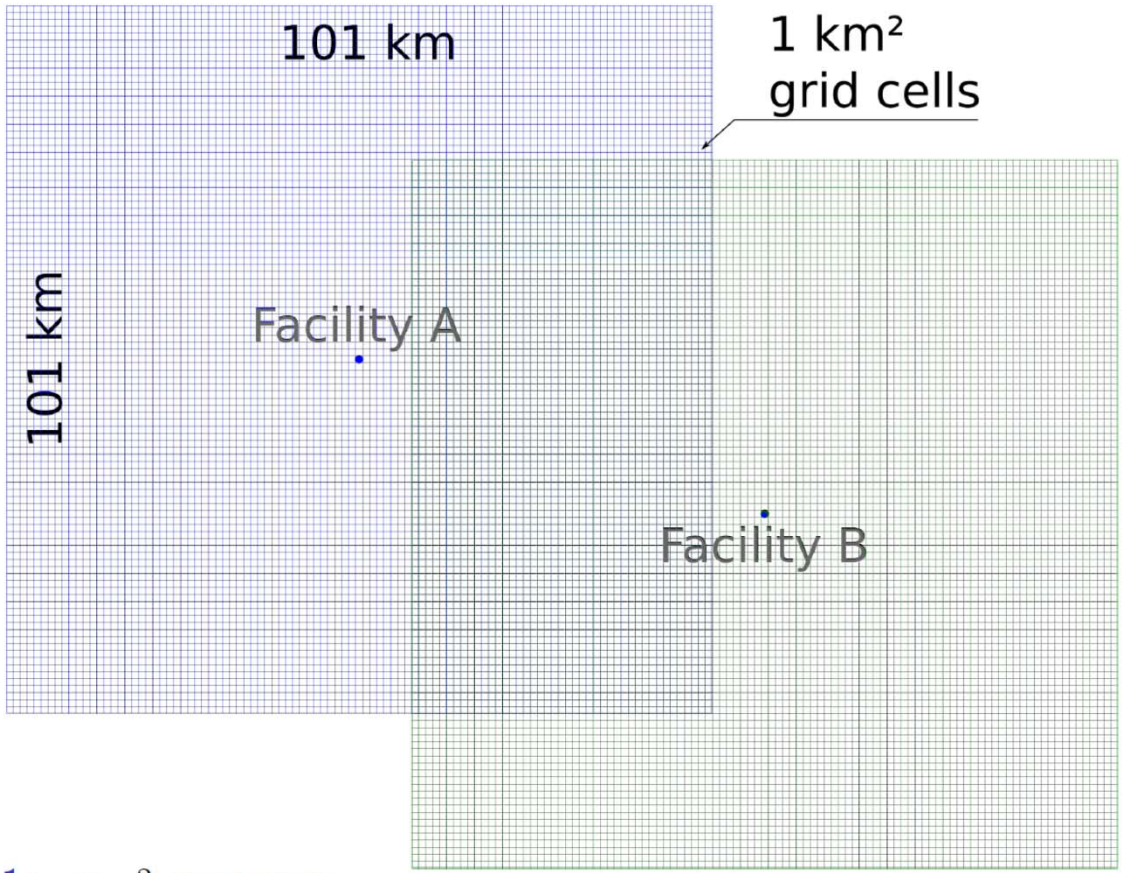


Figure 1: km^2 Grid Cells

RSEI-GM & RSEI public release

- **RSEI-GM (RSEI-Geo Microdata)**: 3 measures at the km^2 level, for each chemical in the TRI database (2000-2002, 2008-2011):
  1. Concentration
  2. Toxicity*Concentration (Hazard based score)
  3. Toxicity*Concentration*Population (Risk Score)

  Not publicly available mostly due to its daunting size.
- **RSEI-public release**: version v2.3.2: Aggregation of RSEI-GM at the facility level for 1996 to 2011. *Publicly available for download on the EPA site.*

"Because EPA developed the RSEI data for the purpose of prioritizing facilities (that is, sources) for enforcement and clean-up, the public-release data are not designed for examining differences among communities (that is, receptors) in terms of their exposure to industrial toxic releases."[2]

RSEI-GM Grid Cells Example

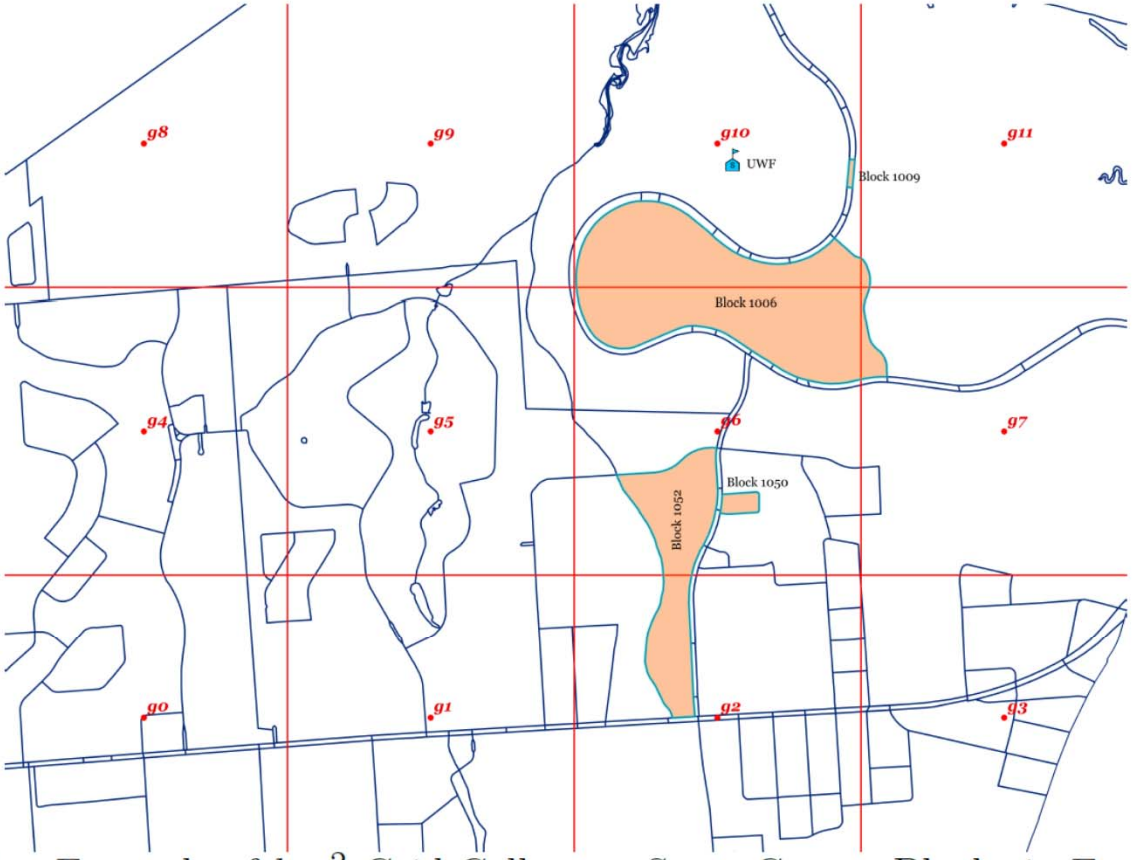


Figure 2: Example of km^2 Grid Cells over Some Census Blocks in Escambia County, FL (2010 Census)

This RSEI-GM study

- Focuses on carcinogenic releases for the Contiguous U.S.
- Rough trend from the RSEI-GM (2000-2002, 2008-2011).
- Zoom-In: County ⇒ Tracts ⇒ Block Groups ⇒ Blocks

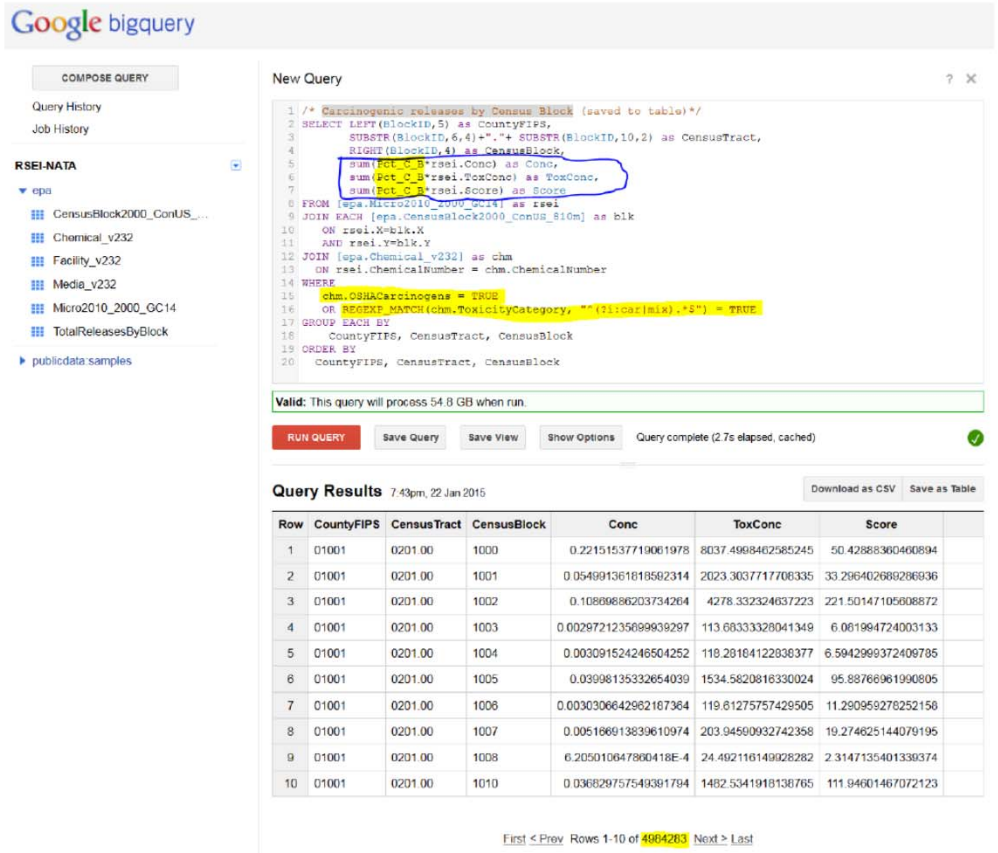
# of counties:	3,003
# of census tracts:	71,901
# of census block groups:[2]	213,821
# of census blocks:[3]	10,203,760

- Primary software used: SAS, ArcGIS, Tableau, SaTScan, Google BigQuery, MS Access

[2] Block groups (BGs) are the next level above census blocks in the geographic hierarchy. A BG consists of all census blocks whose numbers begin with the same digit in a given census tract.[5]

[3] Census blocks, the smallest geographic area for which the Bureau of the Census collects and tabulates decennial census data.[5]

Carcinogenic Releases Google BigQuery



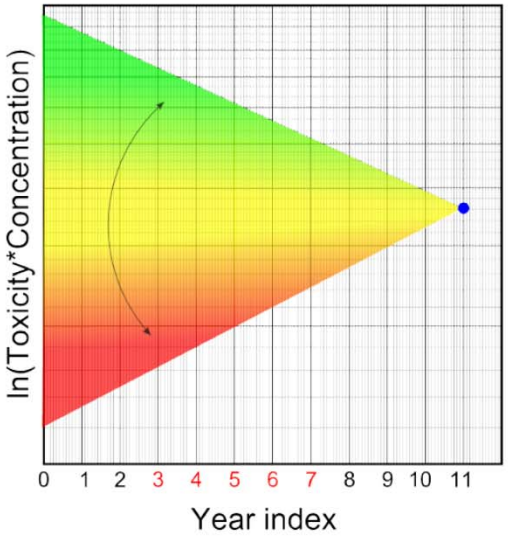
Row	CountyFIPS	CensusTract	CensusBlock	Conc	ToxConc	Score
1	01001	0201.00	1000	0.22151537719061978	8037.4998462585245	50.42888360460894
2	01001	0201.00	1001	0.054991361818592314	2023.3037717708335	33.296402689286936
3	01001	0201.00	1002	0.10869886203734264	4278.332324637223	221.50147105608872
4	01001	0201.00	1003	0.002972123589939297	113.6833332804134 9	6.061994724003133
5	01001	0201.00	1004	0.003091524246504252	118.28164122838377	6.5942999372409785
6	01001	0201.00	1005	0.03998135332654039	1534.5820816330024	95.88768961990805
7	01001	0201.00	1006	0.0030306642962187364	119.61275757429505	11.290959278252158
8	01001	0201.00	1007	0.005166913839610974	203.94590932742358	19.274625144079195
9	01001	0201.00	1008	6.205010647860418E-4	24.492116149928282	2.3147135401339374
10	01001	0201.00	1010	0.036829757549391794	1482.5341918138765	111.94601467072123

Figure 3: Carcinogenic Census Block level aggregation (RSEI-GM 2000)

University of West Florida
R. Amin, T. Charlot

How is pollution in an area changing over time?

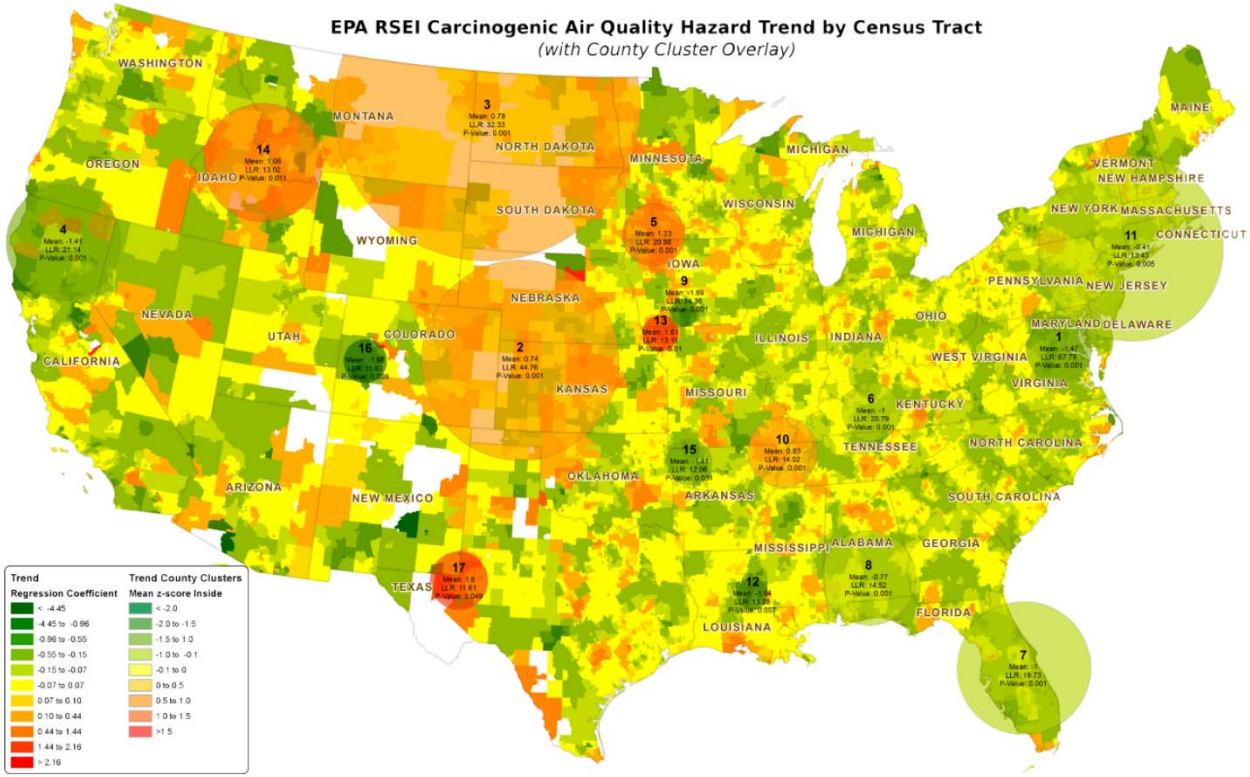
- Performed linear regression analysis on each Census Tract over 7 nonconsecutive years, 72,114 tracts. *(2000-2002 and 2008-2011)*
- The direction of the regression coefficient (the slope) indicates the trend. A negative slope means pollution is generally improving.
- Also independently performed Regression analysis on 3,003 counties.
- Use the z-score of the slopes and performed county level spatial cluster analysis.



University of West Florida
R. Amin, T. Charlot

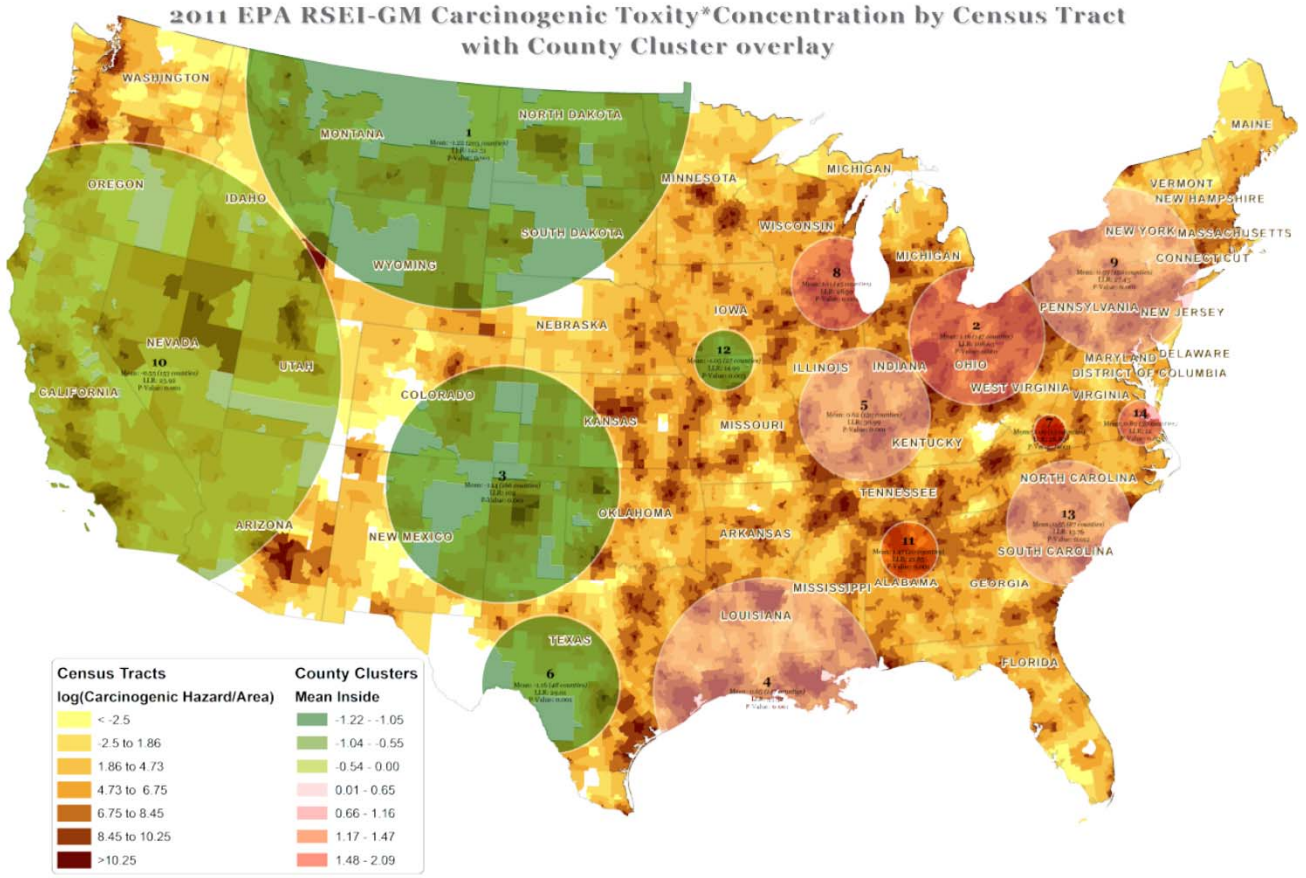
Linear Regression Results

The county clusters are overlaid over the tracts using a transparency for a quick peek at the contents of the cluster.

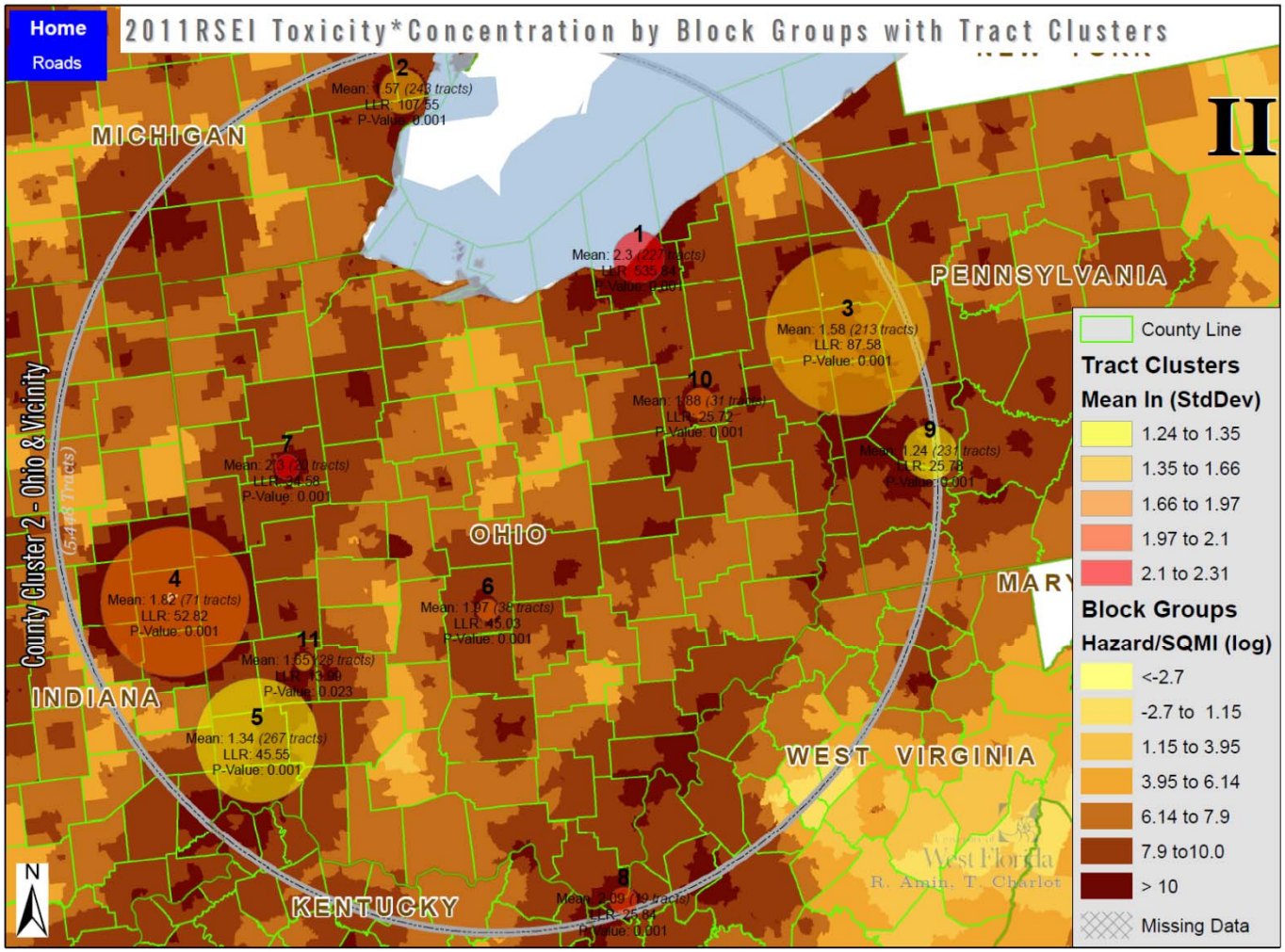


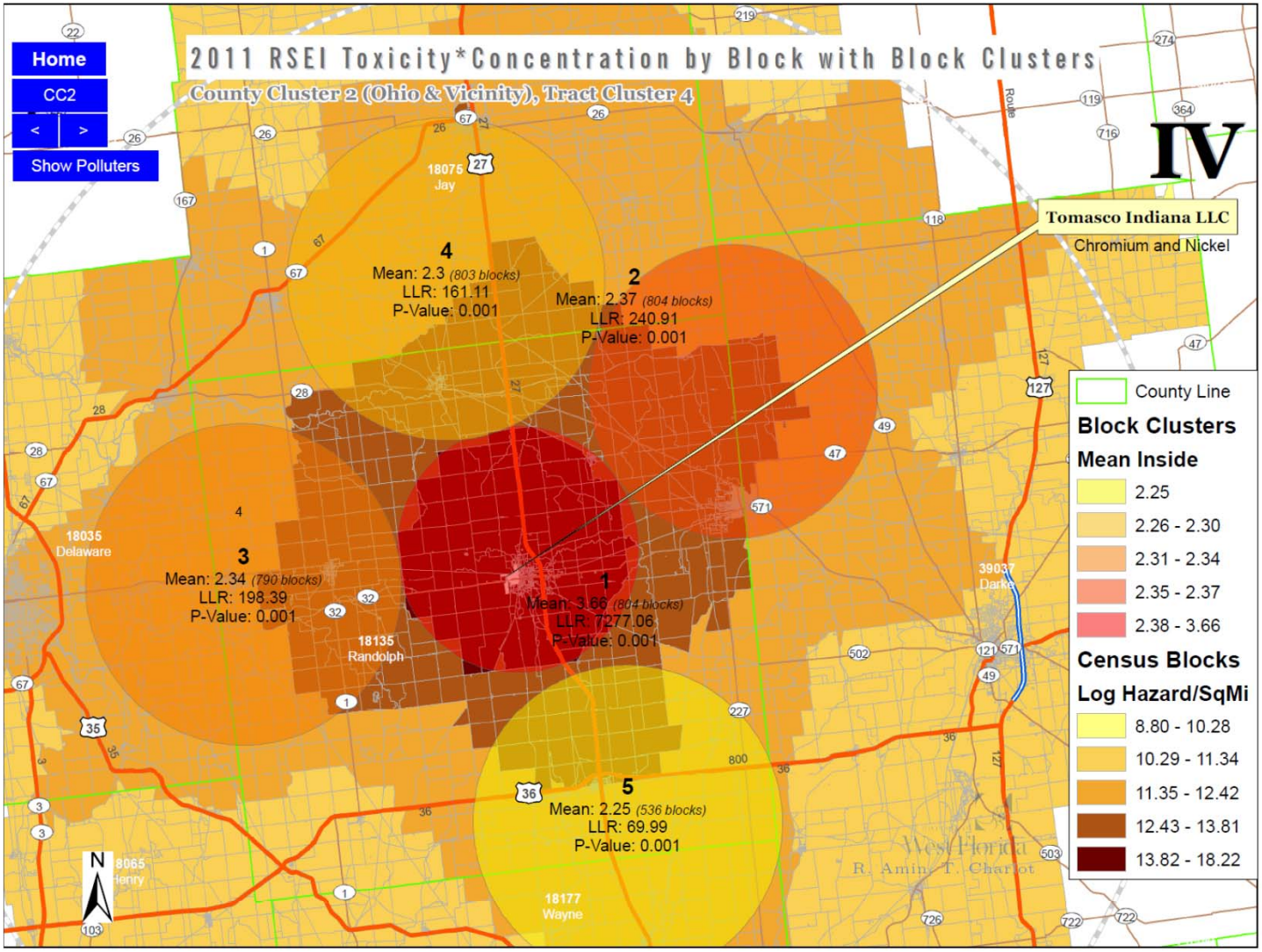
University of West Florida
R. Amin, T. Charlot

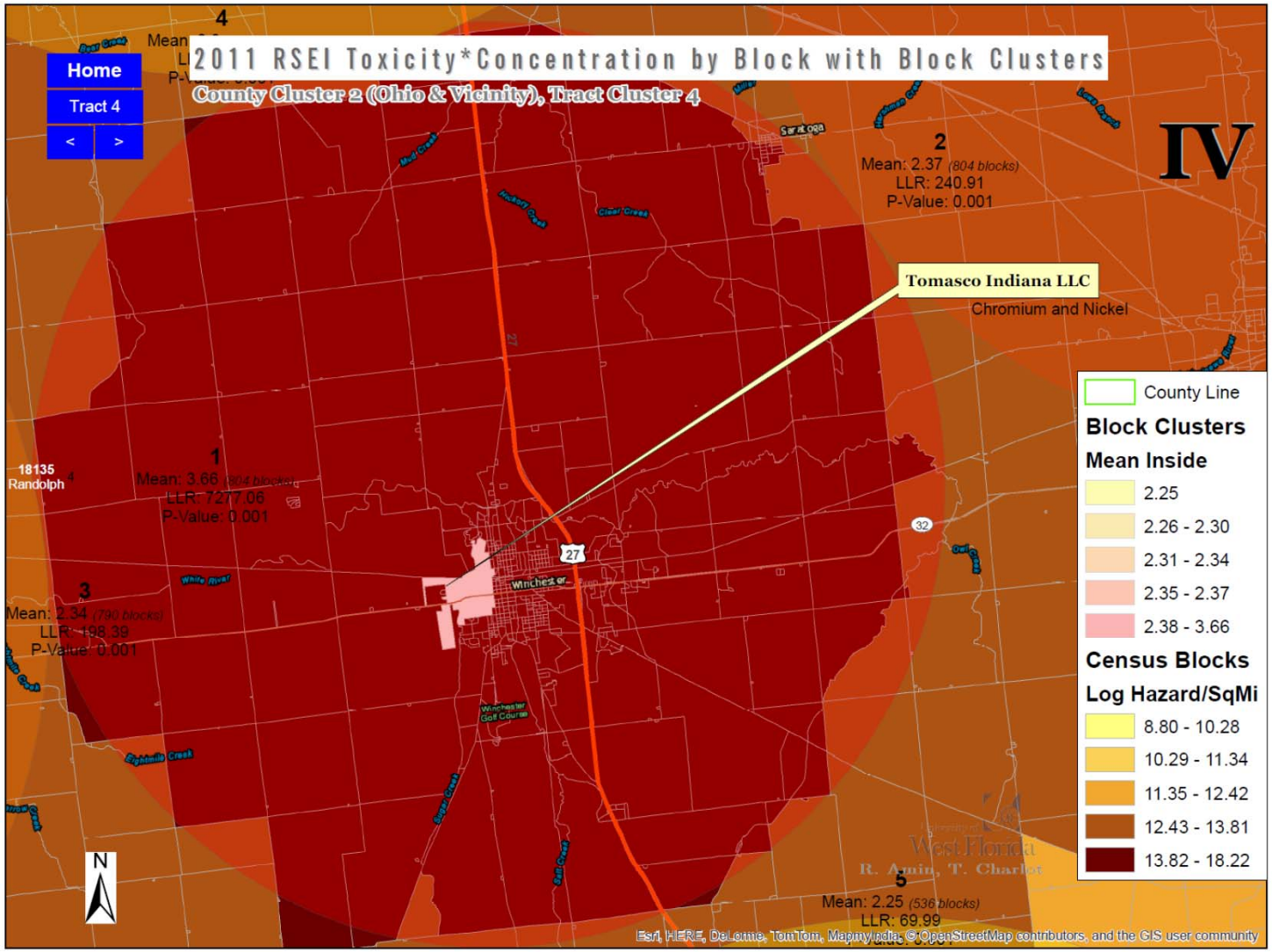
County ⇒ Tracts ⇒ Block Groups ⇒ Blocks



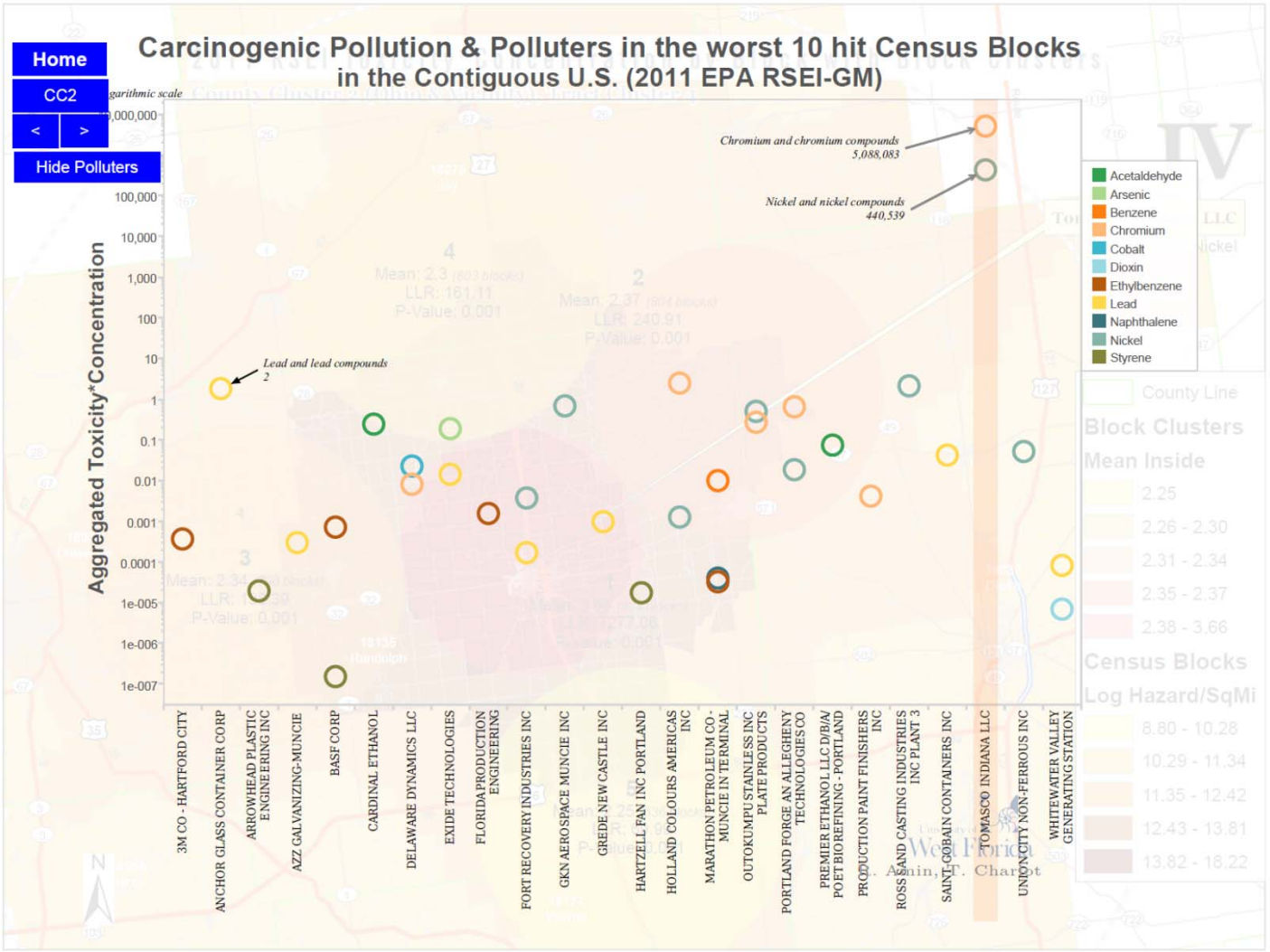
University of West Florida
R. Amin, T. Charlot


Home
Roads
2011RSEI Toxicity*Concentration by Block Groups with Tract Clusters
II
MICHIGAN
PENNSYLVANIA
OHIO
INDIANA
WEST VIRGINIA
KENTUCKY
County Cluster 2 - Ohio & Vicinity
(5,448 Tracts)
2
Mean: 1.57 (243 tracts)
LLR: 107.55
P-Value: 0.001
1
Mean: 2.3 (227 tracts)
LLR: 535.84
P-Value: 0.001
3
Mean: 1.58 (213 tracts)
LLR: 87.58
P-Value: 0.001
10
Mean: 1.88 (31 tracts)
LLR: 25.72
P-Value: 0.001
9
Mean: 1.24 (231 tracts)
LLR: 25.78
P-Value: 0.001
7
Mean: 2.3 (20 tracts)
LLR: 34.58
P-Value: 0.001
4
Mean: 1.82 (71 tracts)
LLR: 52.82
P-Value: 0.001
6
Mean: 1.97 (38 tracts)
LLR: 45.03
P-Value: 0.001
11
Mean: 1.55 (28 tracts)
LLR: 13.09
P-Value: 0.023
5
Mean: 1.34 (267 tracts)
LLR: 45.55
P-Value: 0.001
8
Mean: 2.09 (19 tracts)
LLR: 25.84
P-Value: 0.001
N
County Line
Tract Clusters
Mean ln (StdDev)
1.24 to 1.35
1.35 to 1.66
1.66 to 1.97
1.97 to 2.1
2.1 to 2.31
Block Groups
Hazard/SQMI (log)
<-2.7
-2.7 to 1.15
1.15 to 3.95
3.95 to 6.14
6.14 to 7.9
7.9 to10.0
> 10
Missing Data


Home
CC2
< >
Show Polluters
2011 RSEI Toxicity*Concentration by Block with Block Clusters
County Cluster 2 (Ohio & Vicinity), Tract Cluster 4
IV
Tomasco Indiana LLC
Chromium and Nickel
4
Mean: 2.3 (803 blocks)
LLR: 161.11
P-Value: 0.001
2
Mean: 2.37 (804 blocks)
LLR: 240.91
P-Value: 0.001
3
Mean: 2.34 (790 blocks)
LLR: 198.39
P-Value: 0.001
1
Mean: 3.66 (804 blocks)
LLR: 7277.06
P-Value: 0.001
5
Mean: 2.25 (536 blocks)
LLR: 69.99
P-Value: 0.001
18075 Jay
18035 Delaware
18135 Randolph
18177 Wayne
39037 Darke
18065 Henry
County Line
Block Clusters
Mean Inside
2.25
2.26 - 2.30
2.31 - 2.34
2.35 - 2.37
2.38 - 3.66
Census Blocks
Log Hazard/SqMi
8.80 - 10.28
10.29 - 11.34
11.35 - 12.42
12.43 - 13.81
13.82 - 18.22
N



2011 RSEI Toxicity*Concentration by Block with Block Clusters
County Cluster 2 (Ohio & Vicinity), Tract Cluster 4
Home
Tract 4
IV
Tomasco Indiana LLC
Chromium and Nickel
1
Mean: 3.66 (804 blocks)
LLR: 7277.06
P-Value: 0.001
2
Mean: 2.37 (804 blocks)
LLR: 240.91
P-Value: 0.001
3
Mean: 2.34 (790 blocks)
LLR: 198.39
P-Value: 0.001
4
5
Mean: 2.25 (536 blocks)
LLR: 69.99
18135
Randolph
Winchester
Winchester Golf Course
Saratoga
County Line
Block Clusters
Mean Inside
2.25
2.26 - 2.30
2.31 - 2.34
2.35 - 2.37
2.38 - 3.66
Census Blocks
Log Hazard/SqMi
8.80 - 10.28
10.29 - 11.34
11.35 - 12.42
12.43 - 13.81
13.82 - 18.22
N
Esri, HERE, DeLorme, TomTom, MapmyIndia, © OpenStreetMap contributors, and the GIS user community


Home
CC2
<
>
Hide Polluters
Carcinogenic Pollution & Polluters in the worst 10 hit Census Blocks
in the Contiguous U.S. (2011 EPA RSEI-GM)
garithmic scale
,000,000
100,000
10,000
1,000
100
10
1
0.1
0.01
0.001
0.0001
1e-005
1e-006
1e-007
Aggregated Toxicity*Concentration
Chromium and chromium compounds
5,088,083
Nickel and nickel compounds
440,539
Lead and lead compounds
2
Acetaldehyde
Arsenic
Benzene
Chromium
Cobalt
Dioxin
Ethylbenzene
Lead
Naphthalene
Nickel
Styrene
County Line
Block Clusters
Mean Inside
2.25
2.26 - 2.30
2.31 - 2.34
2.35 - 2.37
2.38 - 3.66
Census Blocks
Log Hazard/SqMi
8.80 - 10.28
10.29 - 11.34
11.35 - 12.42
12.43 - 13.81
13.82 - 18.22
3M CO - HARTFORD CITY
ANCHOR GLASS CONTAINER CORP
ARROWHEAD PLASTIC ENGINEERING INC
AZZ GALVANIZING-MUNCIE
BASF CORP
CARDINAL ETHANOL
DELAWARE DYNAMICS LLC
EXIDE TECHNOLOGIES
FLORIDA PRODUCTION ENGINEERING
FORT RECOVERY INDUSTRIES INC
GKN AEROSPACE MUNCIE INC
GREDE NEW CASTLE INC
HARTZELL FAN INC PORTLAND
HOLLAND COLOURS AMERICAS INC
MARATHON PETROLEUM CO - MUNCIE IN TERMINAL
OUTOKUMPU STAINLESS INC PLATE PRODUCTS
PORTLAND FORGE AN ALLEGHENY TECHNOLOGIES CO
PREMIER ETHANOL, LLC D/B/A/ POET BIO REFINING - PORTLAND
PRODUCTION PAINT FINISHERS INC
ROSS SAND CASTING INDUSTRIES INC PLANT 3
SAINT-GOBAIN CONTAINERS INC
TOMASCO INDIANA LLC
UNION CITY NON-FERROUS INC
WHITEWATER VALLEY GENERATING STATION