
2010 Emissions Modeling Platform Spatial Surrogate Documentation

Contract No: EP-D-12-044
Work Assignment: 1-05
Version No: 1.0

Prepared for: Rich Mason
U.S. EPA OAQPS
C339-02 USEPA Mailroom
Research Triangle Park, NC 27711

Prepared by: Zachariah Adelman
Institute for the Environment
The University of North Carolina at Chapel Hill
137 E. Franklin St., CB 6116
Chapel Hill, NC 27599-6116

Date due: September 30, 2013

Contents

Acknowledgements	4
1 Introduction.....	4
2 Shapefile Database	7
2.1 U.S. Census-based Surrogates	7
2.1.1 Population and Housing	7
2.1.2 Home Heating Fuels	8
2.2 Transportation Surrogates.....	9
2.2.1 Roadways	9
2.2.2 Rail Length	10
2.2.3 Rail Freight Density	10
2.2.4 Rail Yards.....	11
2.2.5 Highway Exits	11
2.2.6 Major Roads	12
2.2.7 Transportation Terminals	12
2.3 Shipping and Port Surrogates	13
2.3.1 Waterways	13
2.3.2 Marine Ports Point Locations	13
2.3.3 Marine Port Area Emissions.....	14
2.3.4 Great Lakes Tugboat Zones.....	15
2.3.5 Gulf of Mexico Platforms.....	16
2.3.6 Offshore Commercial Shipping.....	16
2.3.7 Midwest Shipping and Inland Waterways.....	17
2.4 FEMA Building Footprints.....	18
2.5 Landcover/Landuse Surrogates	19
2.5.1 Agriculture.....	19
2.5.2 Water and Land	19
2.5.3 Developed Areas	20
2.5.4 Forest Land.....	20
2.6 Oil and Gas Production.....	21
2.7 Other Industrial and Commercial Activities.....	25

2.7.1	Mines	25
2.7.2	Quarries	25
2.7.3	Dry Cleaners	26
2.7.4	Gas Stations	26
3	Known Issues and Future Work.....	27
4	Surrogate Revision History	28
5	REFERENCES.....	29

Acknowledgements

ENVIRON International Corporation was a collaborator on this Work Assignment and worked extensively on the development and documentation of the oil and gas production surrogates. They provided all of the content of Section 2.6 of this memo. ERG, Inc. contributed to this work through the development of the Gulf of Mexico platform and support vessel density Shapefile. The Lake Michigan Air Directors Consortium (LADCO) provided the Midwest shipping lane and Inland Waterways Shapefile. The Georgia Department of Natural Resources provided the rail yard point inventory used for generating the rail yard Shapefile and surrogate. Computer Sciences Corporation (CSC) developed the Great Lakes tugboat zone Shapefile and provided support in the testing and QA/QC of the final surrogate data.

1 Introduction

This memo describes the collection, processing, and development of updated geospatial data for calculating spatial surrogates to be used in the EPA 2008 and 2011 emissions modeling platforms. Spatial surrogates are used to allocate county level emission inventories into the rectangular grid cells used by urban and regional scale air quality models. Spatial surrogates are based on data at resolutions different from county-level, such as census tracts or road locations, and can therefore be used to allocate the emissions more specifically than assuming they are uniformly distributed through out the county. For example, motor vehicle emissions from interstate highways can be placed into the grid cells that intersect the highways themselves, and dry cleaning emissions can be allocated to parts of the county that have higher population, or the parts that have more dry cleaners – if data are available on specific locations of dry cleaners. Spatial surrogates consist of values between 0 and 1.0 that specify the fraction of the county emissions that should be allocated to each grid cell that intersects the county. The spatial surrogate fractions for each county sum to 1.0, except the sum may be less than 1 for some counties that intersect the edge and are not completely contained within the modeling domain. The emissions for a specific county within a grid cell are computed by multiplying the surrogate fraction for that county and grid cell by the total emissions for the county (Eyth et al., 2007). A diagram illustrating how spatial surrogates impact the emissions levels within the grid cells of a county is shown in Figure 1.

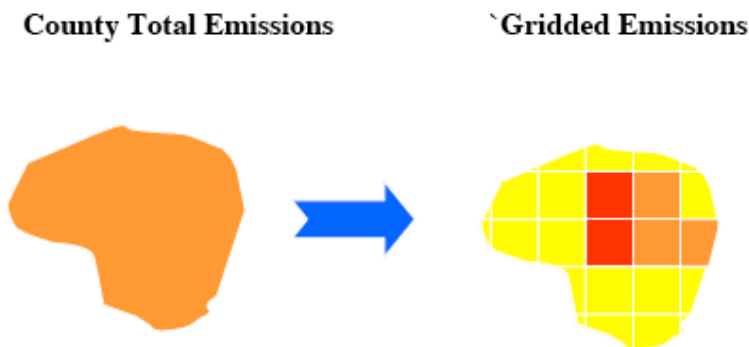


Figure 1. Spatial surrogate schematic

Spatial surrogates have three components:

1. Weight Shapefile – geospatial dataset used for allocating (or weighting) county-level emission inventory sources to a modeling grid. Examples include roadways, population density, or the point locations of marine ports. This memo describes the weight Shapefile updates completed under this Work Assignment.
2. Data Shapefile – geospatial dataset that corresponds to the administrative unit boundaries of the emission inventory. For this Work Assignment, the data Shapefile defines U.S. county boundaries. EPA specified that the year 2000 county boundaries used for the previous version of the national spatial surrogates be used as the data Shapefile for this Work Assignment.
3. Output grid or polygon – definition of the modeling grid(s) or polygon(s) upon which the spatial surrogates map emission inventories. For this Work Assignment, the three output grids include continental U.S. modeling grids at resolutions of 36-km, 12-km, and 4-km.

For air quality modeling studies prior to 2012, EPA used geospatial data primarily from the year 2000 and earlier for the weight Shapefiles that characterize the spatial distribution of non-point anthropogenic emissions sources. In November 2011 the University of North Carolina Institute for the Environment (UNC-IE) initiated updates to the spatial surrogate and Shapefile database used by the EPA Office of Air Quality Planning and Standards for emissions modeling. By July 2012 many of the original surrogates were updated with weight Shapefiles that reflected North American geospatial data collected since the year 2006. Incremental updates through September 2013 targeted specific inventory sectors, such as offshore commercial shipping and on-shore oil and gas production. The data and configurations used to update these surrogates are described here. All of the surrogates described in this memo were generated with the Surrogate Tool¹ of the Spatial Allocator (SA)².

The Fall 2011 release of the 2010 U.S. Census population and housing database provided the basis for several updated surrogates³. Year 2010 roadway data from TIGER and rail/port data from the National Transportation Atlas Database (NTAD)⁴ also supported several surrogate updates. Three entirely new surrogates were created to support the processing of MOVES off-network sources. The Off-Network HD/MD surrogate (251) is designed for MOVES rate per profile (RPP) and rate per vehicle (RPV) medium and heavy-duty vehicles. It's built from a combination of industrial building square footage data and highway exit ramps. The rationale behind this surrogate is that idling and starting/stopping medium and heavy-duty vehicles occur mainly while loading freight in industrial areas and at rest areas and along highway exits ramps. As we were not able to locate a national Shapefile of highway rest areas, we used exit ramps to represent the locations of idling tractor-trailers. The Off-Network LD surrogate (252) is designed for MOVES RPP and RPV light-duty vehicles, including motorcycles. It's built from a combination of commercial building square footage data, residential building square footage data, and local roads. The rationale behind this surrogate is that idling and starting/stopping light

¹ http://www.cmascenter.org/sa-tools/documentation/4.0/html/srgtool/SurrogateToolUserGuide_4_0.htm

² <http://www.cmascenter.org/sa-tools/>

³ <http://www.census.gov/geo/www/tiger/tgrshp2010/tgrshp2010.html>

⁴ http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/index.html

duty vehicles and motorcycles occur mostly in driveways/residential parking areas, parking lots of shopping centers, and along non-residential roadways with street parking spots. The Off-Network Buses surrogate (253) is designed for MOVES RPP and RPV buses. It's built from a combination of bus terminal locations and local roads, with the concept behind this surrogate being that bus terminals and local roads represent the majority of locations that buses are loading/unloading passengers.

Additional surrogate updates included home heating, building square footage, land-used based surrogates, oil and gas production, commercial businesses, and offshore shipping and ports. The 5-year American Community Survey (ACS) release in 2010 included updated data on home heating distributions by Census block⁵. The Federal Emergency Management Administrations (FEMA) HAZUS-MH v2.0 was released in September 2011 and contains updated building square footage data that are applicable to all of the 500-series EPA surrogates⁶. The 30-m resolution, year 2006 National Landcover Database (NLCD) was used to update the land-use based surrogates, such as total land area, forest land, and crops⁷. Year 2008 and 2011 oil and gas production data from IHS, Inc. were aggregated to the continental U.S. (CONUS) modeling grid for the Rocky Mountain and Northeast U.S. states to create twelve new surrogates for different types of oil and gas production. Census block group counts of gas stations and dry cleaners from 2009 were used to create new surrogates for these commercial business categories. Several new datasets of commercial shipping lane activity, tugboat activity, and rail freight density were used to update the surrogates for the aircraft/locomotive/marine inventory sector. Table 1 lists the surrogate classes currently in the EPA emissions modeling platform that have been updated as part of this effort. Table 1 includes the years of the previous and updated geospatial data used to create the different classes of surrogates.

Table 1. Summary of updates to the EPA surrogate database

Shapefile	Surrogates	Current Year	Previous Year
Population and Housing	100-140	2010	2000
Home Heating	150-165	2005-2010	2000
Road and Rail	200-280	2010	2000
Land-use/Land Cover	300-350	2006	1992
Building Square Footage	500-595	2000-2006	1990
Gas Stations and oil & gas	600-699	2008-2011	2000
Shipping	800-820	2010	1999-2003
Other industrial and commercial	850-890	2010	2000

UNC-IE generated spatial surrogates on the CONUS domain at resolutions of 36-km, 12-km, and 4-km. All of the surrogates were generated on a Lambert Conformal Conic projection (lat_0 = 40° north, lon_0 = 97° west, true_lat1 = 33° north, true_lat2 = 45° north) assuming a spherical earth with a radius of 6,370,000 m. The grid definitions of the three surrogate domains are provided in Table 2.

⁵ <http://www.census.gov/acs/www>

⁶ <http://www.fema.gov/plan/prevent/hazus>

⁷ http://www.mrlc.gov/nlcd06_data.php

Table 2. Surrogate domain definitions

Parameter	36km	12km	4km
X-orig	-2,736,000 m	-2,736,000 m	-2,736,000 m
Y-orig	-2,088,000 m	-2,088,000 m	-2,088,000 m
columns	148	444	1332
rows	112	336	1008

UNC-IE used the gapfilling, normalization, and QA tools of the SA to ensure completeness and quality of the surrogates. Normalizing the surrogates eliminates surrogate values that do not equal 1.0, which can result from rounding errors in the surrogate calculations. The QA steps check for surrogates that do not equal 1.0, missing surrogates, and surrogates that are gapfilled. In addition to these tabulated QA summaries, UNC-IE produced spatial plots of the surrogates on each of the three CONUS domains. The QA tables, spatial plots, and a README file describing the surrogates are packaged in a “QA” directory in the surrogate Linux tar archives. The surrogates and accompanying QA data are available from the U.S. EPA CHIEF⁸ website as part of the OAQPS 2008 and 2011 emissions modeling platforms and from the CMAS Center Data Clearinghouse⁹.

The list of GIS Shapefiles used for the updated spatial surrogates and details of how these data were preprocessed for generating surrogates are provided in Section 2. An Excel Spreadsheet (US_SpatialSurrogate_Workbook_v093013.xlsx) that accompanies this document includes the detailed specifications of the Shapefiles and surrogates that were collected and updated under this Work Assignment.

2 Shapefile Database

2.1 U.S. Census-based Surrogates

2.1.1 Population and Housing

The 2010 TIGER/Line database contains U.S. Census population and housing unit counts at the Census block level for each state. UNC-IE downloaded the entire database, merged the state-level data into a national Shapefile, and projected the data to a U.S. national Lambert Conformal projection on a normal sphere. UNC-IE calculated urban and rural areas using Census block groups. Urban was defined as Census block groups that have a population density of at least 1,000 people per square mile and everything else was defined as rural. In addition, Census block groups are classified based on the number of housing units per square kilometers (housing unit density [HUD]) and the attribute AREA_CODE is assigned to census block groups with the following conditions:

1. Urban: HUD $\geq$ 1000
2. Suburban: HUD $\geq$ 125 and HUD $<$ 1000
3. Exurban: HUD $\geq$ 6 and HUD $<$ 125

⁸ <http://www.epa.gov/ttnchie1/>

⁹ <http://www.cmascenter.org/download/data.cfm>

4. Rural: HUD < 6

The attributes in the file used to compute surrogates include:

- POP2010 – 2010 population
- HU2010 – 2010 housing units
- URBAN – urban population in 2010
- RURAL – rural population in 2010
- HUCH1000 – housing unit changes from 2000 to 2010
- AREA_CODE – census block group codes; 1=urban, 2=suburban, 3=exurban, 4=rural

Shapefile Name: pophu_bg2010

Shapefile Type: Polygon

Year: 2010

Attributes: STFID, FIRST_STAT, FIRST_COUN, FIRST_TRAC, FIRST_BLOC, POP2010, HU2010, POP2000, HU2000, HUCH1000, POPCH1000, BG_AREA, URBAN10, POPD10, URBAN, RURAL, AREA_CODE

Projection: Lambert Conformal Conic (X₀=-97.0, Y₁=33.0, Y₂=45.0, Y₀=40.0)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 100, 110, 120, 130, 131, 132, 133, 134, 137, 140, 165, 400

2.1.2 Home Heating Fuels

The American Community Survey (ACS) is a U.S. Census project that collects yearly demographic and housing information from randomly selected households throughout the U.S. The data are aggregated in 5 and 10-year increments to provide Census tract estimates for the statistics collected during the survey. Home heating fuel type from the ACS was used to develop spatial surrogates for home heating sources (i.e. residential wood combustion) in the nonpoint inventory. The ACS 5-year 2010 survey results that were released in 2011 represent data collected from 2006-2010. These data were used to create home heating surrogates.

As a review of the 5-year ACS data for 2010 showed that the number of housing units per Census tract in the ACS are always lower than the number of housing units reported in the 2010 Census. This trend indicates that the ACS only estimates demographic and housing statistics for a random sampling of households in each Census tract and does not represent the entire distribution of households. In order to represent the spatial distribution of home heating sources accurately, the ACS was used to calculate a distribution of heating sources for each Census tract and then these distributions were applied to the 2010 Census Housing Unit Shapefile to construct a home heating Shapefile that is consistent with the number of census tract housing units in the 2010 Census database. For example, if the ACS reports that in Census tract X 75% of the households use coal for heating and 25% of the households use heating fuel, the 2010 U.S. Census estimate of the total number of households in tract X were multiplied by 0.75 and 0.25 to estimate the distribution of coal and heating fuel use, respectively, for tract X.

The attributes in the file used to compute surrogates include:

- UTIL_GAS - number of housing units using Utility Gas for primary heating

- WOOD - number of housing units using Wood for primary heating
- FUEL_OIL - number of housing units using Fuel Oil for primary heating
- COAL - number of housing units using Coal for primary heating
- LP_GAS - number of housing units using Bottled Gas for primary heating

Shapefile Name: heating_fuels_acs0510_c2010

Shapefile Type: Polygon

Year: 2011

Attributes: UTIL_GAS, WOOD, FUEL_OIL, LP_GAS

Projection: Lambert Conformal Conic (X₀=-97.0, Y₁=33.0, Y₂=45.0, Y₀=40.0)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 150, 160, 165, 170, 180, 190

2.2 Transportation Surrogates

2.2.1 Roadways

UNC-IE merged state-level 2010 TIGER/Line Shapefile data with population density data to create a national file with urban and rural roadways. UNC-IE projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere. The TIGER/Line MTFCC codes S1100 and S1200 were used to define primary and secondary roads, respectively. Urban and rural roadway classifications were calculated by UNC-IE using Census block group population densities (see Section 2.1.1) overlaid onto TIGER roads.

The attributes in the file used to compute surrogates include:

- RDTYPE = road types; 1=urban primary, 2=rural primary, 3=urban secondary, 4=rural secondary

Shapefile Name: rd_ps_tiger2010

Shapefile Type: Line

Year: 2010

Attributes: STATEFP, LINEARID, FULLNAME, RTTYP, MTFCC, STFID, URBAN10, RDTYPE

Projection: Lambert Conformal Conic (X₀=-97.0, Y₁=33.0, Y₂=45.0, Y₀=40.0)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 200, 210, 220, 230, 240, 250

2.2.2 Rail Length

UNC-IE merged state-level 2010 TIGER/Line Shapefile data to create a national railway network file. The TIGER/Line MTFCC codes R1052, R1051, and R1011 were used to define the different classes of rail lines. UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere.

The attributes in the file used to compute surrogates include:

- RRTYPE – rail classes; 1=class 1; 2=class 2; 3=class 3

Shapefile Name: rail_tiger2010

Shapefile Type: Line

Year: 2010

Attributes: LINEARID, FULLNAME, MTFCC, RRTYPE

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 260, 270, 280

2.2.3 Rail Freight Density

UNC-IE used the 2012 National Transportation Atlas Database (NTAD) freight densities to weight the line Shapefile of rail activity locations. The NTAD RAILTYPE codes were used to define the different classes of rail lines. UNC-IE projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere.

The attributes in the file used to compute surrogates include:

- RAILTYPE – rail classes; 1=class 1,2,3 rail; 2=Amtrak; 3=Commuter rail

Shapefile Name: rail_lines_ntad2012

Shapefile Type: Line

Year: 2012

Attributes FRAARCID, MILES, STATEAB, STATEFIPS, CNTYFIPS, STCNTYFIPS, FRAREGIONS, RROWNER1, RROWNER2, RROWNER3, TRKRGHTS1, TRKRGHTS2, TRKRGHTS3, TRKRGHTS4, TRKRGHTS5, TRKRGHTS6, TRKRGHTS7, TRKRGHTS8, TRKRGHTS9, STRACNET, SIGSYS, TRACKS, SUBDIV, FRFRANODE, TOFRANODE, NET, PASSNGR, YARDS, INT_TYPE, DEN09CODE, VERSION, MEDDENS, RAILTYPE

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 261, 271, 272, 273

2.2.4 Rail Yards

Michelle Bergin and Byeong Kim of GA DNR provided a spreadsheet of rail yard locations and emissions collected as part of the Eastern Regional Technical Advisory Committee (ERTAC) railway inventory development and improvement effort (Bergin et al., 2009). UNC-IE converted this spreadsheet to a point Shapefile for generating a rail yard surrogate for the U.S. The surrogate is based on the sum of the NO_x and PM_{2.5} emissions at the rail yard locations in the ERTAC rail yard inventory.

The attributes in the file used to compute surrogates include:

- NOXPMTOT – total of NO_x and PM_{2.5} emissions at each rail yard

Shapefile Name: ERTAC_railyard_WRF

Shapefile Type: Point

Year: 2012

Attributes OBJECTID, OWNER1, OWNER2, OWNER3, STATE, FIPS, YARDNAME, LON, LAT, NOX, PM25, NOXPMTOT,

Projection: Lambert Conformal Conic (X₀=-97.0, Y₁=33.0, Y₂=45.0, Y₀=40.0)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 275

2.2.5 Highway Exits

This point Shapefile of highway exit ramp locations was obtained from the ESRI Data and Maps 2010 database. UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere. As these data are proprietary, this Shapefile cannot be redistributed.

The attributes in the file used to compute surrogates include:

- NONE – default to point locations

Shapefile Name: exits_lcc

Shapefile Type: Point

Year: 2010

Attributes: EXIT_ID, FROM_NAME, EXIT_NUM, TO_NAME, COUNTRY

Projection: Lambert Conformal Conic (X₀=-97.0, Y₁=33.0, Y₂=45.0, Y₀=40.0)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 201, 251

2.2.6 Major Roads

2010 U.S. Major roads, including interstates, freeways, U.S. and state highways, major streets and roads, primary, secondary, and local roads, access ramps, ferry crossing, and other major thoroughfares within the U.S. This Shapefile was used to define the locations of local and minor roadways for allocating off-network light-duty vehicle MOVES emissions to modeling grids. This line Shapefile of major roads was obtained from the ESRI Data and Maps 2010 database. UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere. As these data are proprietary, this Shapefile cannot be redistributed

The attributes in the file used to compute surrogates include:

- FDC – road classes; 4=secondary road; 5=local connecting road

Shapefile Name: mjrds_lcc

Shapefile Type: Line

Year: 2010

Attributes: PREFIX, NAME, TYPE, SUFFIX, FDC, SHIELD, HWY_NUM, OID

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 252

2.2.7 Transportation Terminals

2010 U.S. and Canada locations for transportation terminals such as bus terminals, train stations, marine and ferry terminals, and other significant transportation nodes. This Shapefile was used to define the locations of bus stations for allocating off-network MOVES emissions to modeling grids. This point Shapefile of transportation terminals was obtained from the ESRI Data and Maps 2010 database. UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere. As these data are proprietary, this Shapefile cannot be redistributed.

The attributes in the file used to compute surrogates include:

- FCC – feature class codes; D53=bus terminals

Shapefile Name: transterm

Shapefile Type: Point

Year: 2010

Attributes: NAME, FCC

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 203, 253

2.3 Shipping and Port Surrogates

2.3.1 Waterways

2011 National Transportation Atlas Database (NTAD) data for waterway lengths. UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere.

The attributes in the file used to compute surrogates include:

- LENGTH – navigable waterway lengths

Shapefile Name: waterway_ntad2011

Shapefile Type: Line

Year: 2011

Attributes: LENGTH, FEATUREID, ANODE, BNODE, LINKNAME, RIVERNAME, AMILE, BMILE, LENGTH1, LENTH_SRC, LINKTYPE, CTRL_DEPTH, WATERWAY, GEO_CLASS, FUNC_CLASS, WTWY_TYPE, CHART_ID, NUM_PAIRS, WHO_MOD, DATE_MOD, HEADING, STATE, FIPS, FIPS2, NONUS, VERSION

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 807

2.3.2 Marine Ports Point Locations

2010 NTAD data for marine ports. UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere.

The attributes in the file used to compute surrogates include:

- NONE – default to point locations

Shapefile Name: ports_ntad2010

Shapefile Type: point

Year: 2010

Attributes: NAME, LOCATION, ADDRESS, COUNTY, COUNTYFIPS, TOWN, STATE, WATERWAY, PORT_NAME, MILE, BANK, LATDEC, LONG_DEC, OPERATOR1, OWNER, PURPOSE, RWY_CONN, PORTSERIES, SEQNO, LOC_CD, PWDNO, OLDPWD, DOCKCD, NDCCODE, CNGDST, COMMODITY1, COMMODITY2, COMMODITY3, COMMODITY4, REMARKS, DATUM, DEPTH1, DEPTH1A, DEPTH2, DEPTH2A, DEPTH3, DEPTH3A, TOTBERTH1, TOTBERTH2, TOTBERTH3, YEAR, MAPNO, FIRSTNAME, LASTNAME, PHONE, FAX, STFIPS, VERSION

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 800

2.3.3 Marine Port Area Emissions

2012 marine port areas weighted by NEI2011 NO_x emissions. UNC-IE joined a Shapefile of marine port areas with NEI2011 NO_x emissions for SCC 2280002100. UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere.

The attributes in the file used to compute surrogates include:

- ann_value – annual NO_x emissions for port vessels from the NEI2011

Shapefile Name: Ports_081412_NEI2011_NO_x_WRF

Shapefile Type: polygon

Year: 2012

Attributes: NAME,FIPS_1,Area_sqmi,ShapeID,PortID,OID_,country_cd,region_cd,tribal_cod,census_tra,shape_id,scc,emis_type,poll,ann_value,ann_pct_re,control_id,control_me,current_co,cumulative,projection,reg_codes,calc_metho,calc_year,date_updat,data_set_i,jan_value,feb_value,mar_value,apr_value,may_value,jun_value,jul_value,aug_value,sep_value,oct_value,nov_value,dec_value,jan_pctred,feb_pctred,mar_pctred,apr_pctred,may_pctred,jun_pctred,jul_pctred,aug_pctred,sep_pctred,oct_pctred,nov_pctred,dec_pctred,comment,SHPID

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 820

2.3.4 Great Lakes Tugboat Zones



Figure 2. 2011 NEI Great Lakes tugboat zones

U.S. EPA defined a tugboat operating zone in the Great Lakes as 0.5-2.0 miles offshore (Figure 2). A Shapefile that defines the tugboat zone by county was developed for use in creating a spatial surrogate for emissions in the Great Lakes.

The attributes in the file used to compute surrogates include:

- AREA – area of tugboat zone by county

Shapefile Name: TUGvessel_SURROGATE_GREATLAKES

Shapefile Type: Polygon

Year: 2012

Attributes: FIPS, STATE, COUNTY, NAME, ST, STATECTY, AREA,

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 808

2.3.5 Gulf of Mexico Platforms

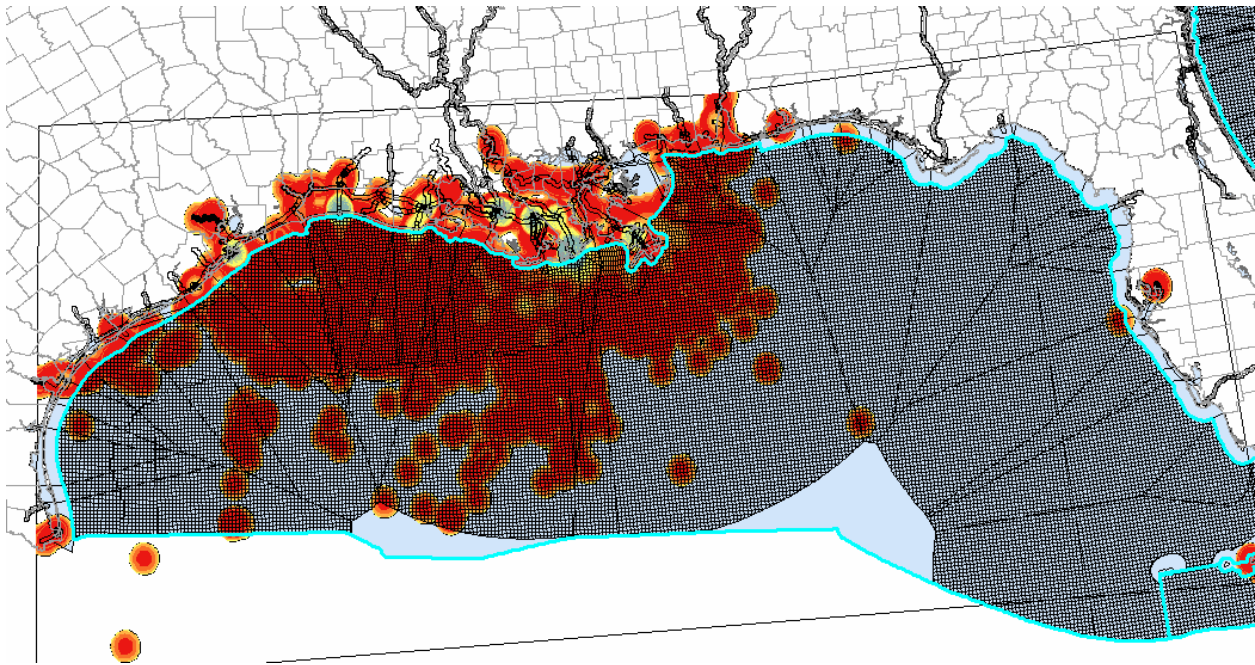


Figure 3. Gulf of Mexico oil and gas platform support vessel densities

The locations of class 1 and 2 commercial shipping vessels in the Gulf of Mexico were tied to the density of shipping activity surrounding oil and gas platforms (Figure 3). A Shapefile that defines the fraction of total Gulf activity density around the energy development platforms was used to create a Gulf Shipping Lanes surrogate.

The attributes in the file used to compute surrogates include:

- Fraction – fraction of Gulf-wide activity associated with individual oil and gas platforms

Shapefile Name: GulfofMexico_SupportVessels_Density_WRF

Shapefile Type: Polygon

Year: 2010

Attributes: GRIDCODE, Area, Fraction

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 803, 802

2.3.6 Offshore Commercial Shipping

Offshore shipping lane areas for the Pacific, Atlantic, Gulf of Mexico, and Great Lakes weighted by NEI2011 NO_x emissions. UNC-IE joined a Shapefile of offshore commercial shipping lane

areas with NEI2011 NO_x emissions for SCC 2280002200. UNC projected the data to a U.S. national Lambert Conformal projection on a normal sphere

The attributes in the file used to compute surrogates include:

- ann_value— annual NO_x emissions for C1/C2 vessels from the NEI2011

Shapefile Name: ShippingLanes_112812_NEI2011_NO_x_WRF

Shapefile Type: Polygon

Year: 2012

Attributes: FIPS,Area_sqmi,ShapeID,OID_,country_cd,region_cd,tribal_cod,census_tra, shape_id, scc,emis_type,poll,ann_value,ann_pct_re,control_id,control_me,current_co, cumulative, projection,reg_codes,calc_metho,calc_year,date_updat,data_set_i,jan_value, feb_value,mar_value,apr_value,may_value,jun_value,jul_value,aug_value,sep_value, oct_value,nov_value,dec_value,jan_pctred,feb_pctred,mar_pctred,apr_pctred,may_pctred, jun_pctred,jul_pctred,aug_pctred,sep_pctred,oct_pctred,nov_pctred,dec_pctred,comment,SHPID

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 805, 806, 802

2.3.7 Midwest Shipping and Inland Waterways

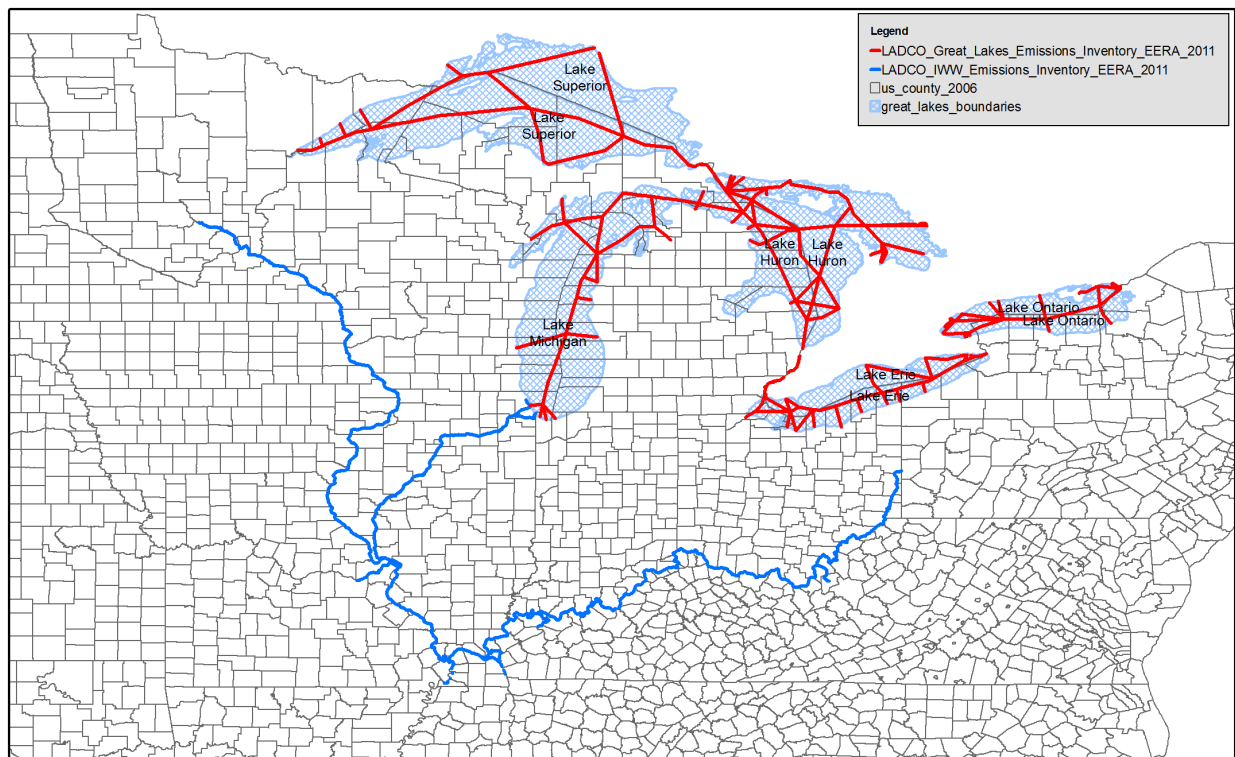


Figure 4. Great Lakes shipping lane and inland waterway inventory

LADCO provided a 2011 shipping lane and inland waterway emission inventory for the Great Lakes region (Figure 4). UNC projected the data to a U.S. national Lambert Conformal Conic projection on a normal sphere and used these data to generate a surrogate for the regional shipping lane emissions.

The attributes in the file used to compute surrogates include:

- G_nonCO2– annual non-CO2 emissions for C1/C2 vessels from the LADCO 2011 inventory

Shapefile Name: GreatLakes_IWW_EERA_2011_WRF

Shapefile Type: Line

Year: 2011

Attributes: NWN_ID,G_CO2GAL,G_NOXGAL,G_SOXGAL,G_HCGAL,G_PMGAL,
G_BCGAL, G_POMGAL,G_COGAL,seg_km,G_nonCO2,GROUP,Seg_Miles,G_CO2,
G_NOX,G_SOX,G_HC,G_PM,G_BC,G_POM,G_CO,

Projection: Lambert Conformal Conic (X₀=-97.0, Y₁=33.0, Y₂=45.0, Y₀=40.0)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 812, 802

2.4 FEMA Building Footprints

HAZUS is a nationally standardized methodology that contains models for estimating potential losses from natural disasters. A key component of the HAZUS database is information on building locations, sizes, and classifications. These geospatial data on buildings are useful for allocating inventory data to modeling grids because they contain details on the locations and sizes of different building types. The Federal Emergency Management Agency HAZUS-MH version 2 was released in September 2011 and contains square footage data for different types of buildings throughout the U.S. The building square footage data are used to identify building classifications (i.e. commercial, residential, industrial, institutional) for allocating non-point inventory sources to modeling grids. UNC merged Census-block level data in state Shapefiles into a national Shapefile for select database attributes. UNC used 2002 Census block boundaries when merging these data to maintain consistency with the previous spatial surrogates used by EPA.

The attributes in the file used to compute surrogates include:

- COM# - commercial building square footage
- IND# - industrial building square footage
- RES# - residential building square footage
- EDU# - educational building square footage

- REL# - religious building square footage
- GOV# - government building square footage

Shapefile Name: fema_bsf_2002bnd

Shapefile Type: Polygon

Year: 2011

Attributes: COM#, IND#, RES#, EDU#, REL#, GOV#

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 500, 505, 510, 512, 515, 520, 525, 527, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 596

2.5 Landcover/Landuse Surrogates

2.5.1 Agriculture

2006 National Land Cover Database (NLCD) land cover data was used to create this Shapefile. UNC-IE re-gridded the original 30-meter resolution data to 250-meter resolution using the dominant land use class. UNC-IE extracted agricultural classes 81 and 82 into a 250-meter raster file and then converted the raster file to a polygon Shapefile.

The attributes in the file used to compute surrogates include:

- GRID_CODE – 81 for Pasture or Hay class and 82 for Crop class

Shapefile Name: nlcd2006_80s_agri

Shapefile Type: Polygon

Year: 2006

Attributes: GRID_CODE

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 310, 318, 319

2.5.2 Water and Land

2006 NLCD land cover data was used to create this Shapefile. UNC-IE re-gridded the original 30-meter resolution data to 250-meter resolution using the dominant land use class. UNC-IE extracted water (class 11) and land (class 0) into a 250-meter raster file and then converted the raster file to a polygon Shapefile.

The attributes in the file used to compute surrogates include:

- GRID_CODE – 11=water; 0=land

Shapefile Name: nlcd2006_10s_waterland

Shapefile Type: Polygon

Year: 2006

Attributes: GRID_CODE

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 340, 346, 350

2.5.3 Developed Areas

2006 NLCD land cover data was used to create this Shapefile. UNC-IE re-gridded the original 30-meter resolution data to 250-meter resolution using the dominant land use class. UNC-IE extracted developed land use classes (21 to 24) into a 250-meter raster file and then converted the raster file to a polygon Shapefile.

The attributes in the file used to compute surrogates include:

- GRID_CODE – 21=Open Space; 22=Developed, Low Intensity; 23=Developed, Medium Intensity; 24=Developed, High Intensity

Shapefile Name: nlcd2006_20s_developed

Type: Polygon

Year: 2006

Attributes: GRID_CODE

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 300, 301, 302, 303

2.5.4 Forest Land

2006 NLCD land cover data was used to create this Shapefile. UNC-IE re-gridded the original 30-meter resolution data to 250-meter resolution using the dominant land use class. UNC-IE extracted the forest land use classes (41,42,43, and 90) into a 250-meter raster file and then converted the raster file to a polygon Shapefile.

The attributes in the file used to compute surrogates include:

- GRID_CODE – 40=Forest land

Shapefile Name: nlcd2006_40_forest

Type: Polygon

Year: 2006

Attributes: GRID_CODE

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 320

2.6 Oil and Gas Production

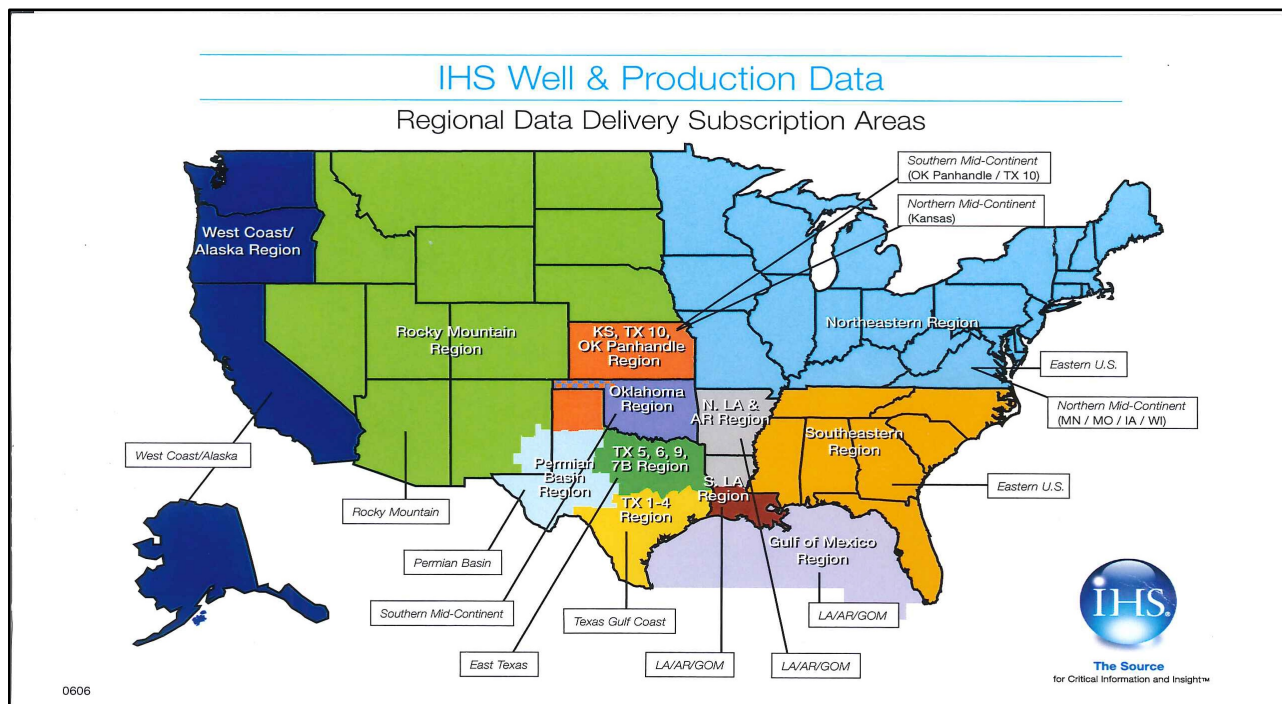


Figure 5. IHS database regions

2008 O&G activities Shapefile for the Rocky Mountain Region

ENVIRON developed a region-wide oil and gas (O&G) emissions inventory for 2008 in the western US for the West-wide Jump Start Air Quality Modeling Study (WestJumpAQMS) managed by the Western Governors' Association. For the WestJumpAQMS, ENVIRON also developed gridding surrogates based on 2008 O&G activities (production and drilling) obtained from the IHS database for use in the SMOKE modeling system. The Rocky Mountain region (Figure 5) Shapefile was based on this existing O&G activities dataset.

ENVIRON developed the Shapefile by aggregating 2008 O&G activities to a 2-km grid. A threshold of three or more wells per grid-cell was applied (i.e. grid-cells having three or more wells were kept in the Shapefile and the remaining cells were dropped). The aggregation was performed to mask any proprietary information contained in the original Shapefiles. As a

consequence, approximately 6% of oil production and 8% of gas production were dropped from the Shapefiles. Figure 6 presents a display of O&G activities dropped due to the threshold (red grid-cells) and O&G activities in the Shapefile (green grid-cells).

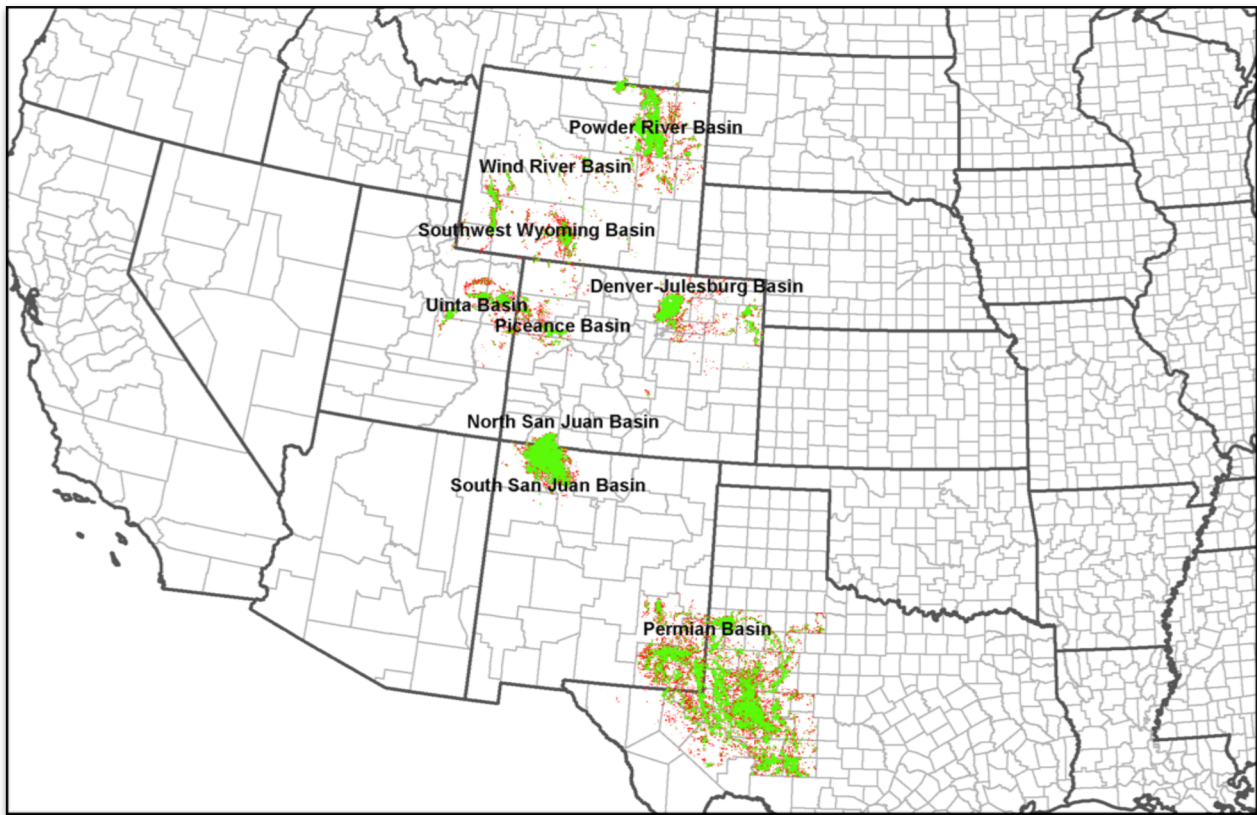


Figure 6. 2008 O&G activities dropped due to threshold for the Rocky Mountain region

2011 O&G activities shapefile for the Northeastern Region

A similar Shapefile was prepared for the Northeastern Region as defined by the IHS database in Figure 5. The Shapefile format is identical to the Rocky Mountain region Shapefile described above. The Shapefile is based on the 2011 O&G activity data from the IHS database. This was the first use by ENVIRON of the Northeastern IHS dataset. The quality of data in the Northeastern IHS dataset differed in some respects from the Rocky Mountain region:

- Gas production data was not available for the following states due to state reporting requirements: IL, IN, KY.
- Oil production data for KY is based on an incomplete dataset. IHS must obtain data from refiners due to lack of state reporting requirements. Comparison of IHS to the US Department of Energy Energy Information Administration (EIA) data indicated that the IHS data included only approximately 40% of the total KY oil production in 2011.
- Except for the above notes, comparisons of the compiled 2011 state-level IHS data to EIA state-level estimates indicated gas production estimates within 7% of EIA estimates for gas production and 15% for oil production. These differences may be due to IHS incorporation of more recent reported volumes than the EIA data.

- ENVIRON found that a number of wells in the IHS database are missing lat-long coordinates. Gap-filling of lat-long coordinates for some wells was possible based on their Township, Range and Section (TRS) (Public Land Survey System) information.
- A small number of IHS records have no available location information and so cannot be translated into surrogates. The following is a listing of the fraction of wells for which location is not included in the Northeastern region:
 - Oil production: 12.8%
 - Gas production: 0.3%
 - Active well count: 1.2%

An analysis was conducted of the aggregation of production data to 2 km grid cells, applying the same threshold of three or more wells per grid cell. The results indicated that approximately 24% of oil production and approximately 13% of gas production would be dropped if this threshold were used. This does not represent a reasonably small fraction of the production that would be dropped; therefore a similar analysis for 3 km and 4 km grid resolutions was performed. Table 4 below shows the results of this analysis.

Table 3. Analysis results for 2011 O&G activities dropped by grid resolution for the Northeastern region

	All Wells with location		Retained Wells *		Difference		% Difference	
Grid resolution	OIL (BBL)	GAS (MCF)	OIL (BBL)	GAS (MCF)	OIL (BBL)	GAS (MCF)	OIL (BBL)	GAS (MCF)
2km Resolution	25,873,260	2,003,377,274	19,567,620	1,745,387,939	6,305,640	257,989,335	24.37%	12.88%
3km Resolution	25,873,260	2,003,377,274	22,062,519	1,845,365,224	3,810,741	158,012,050	14.73%	7.89%
4km Resolution	25,873,260	2,003,377,274	22,981,769	1,914,400,996	2,891,491	88,976,278	11.18%	4.44%

In consultation with EPA, it was decided to develop the Shapefile by aggregating 2011 O&G activities to the 4 km grid-cell. The 4 km grid was defined such that it aligns with the CONUS 4 km domain. A threshold of three or more wells per grid-cell was used, similar to the Rocky Mountain region. Using these criteria, approximately 11% of oil production and approximately 5% of gas production were dropped. Figure 7 shows Northeastern region O&G activities dropped due to the threshold (red grid-cells) and O&G activities in the Shapefile (green grid cells).

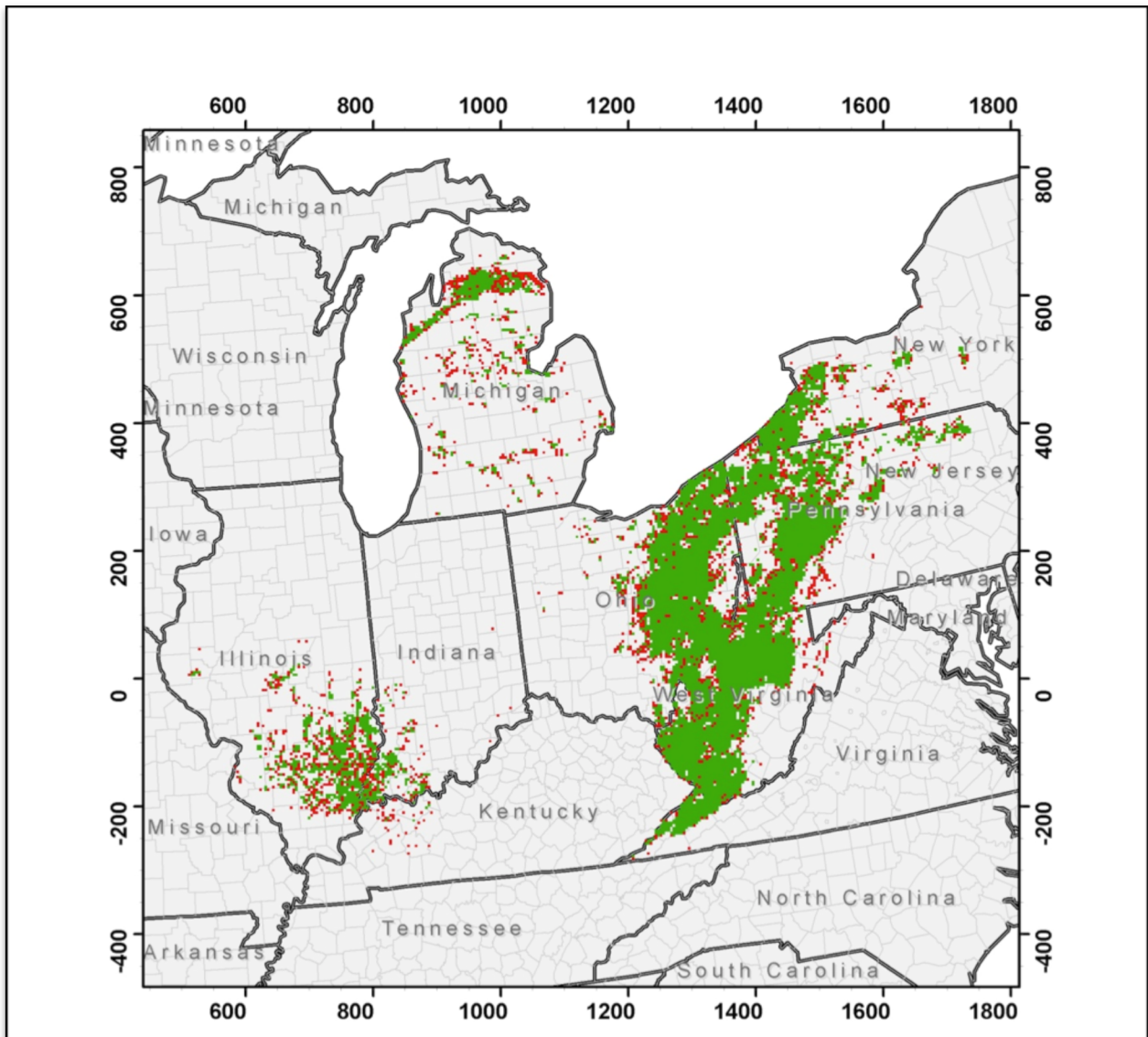


Figure 7. 2011 O&G activities dropped due to threshold for the Northeastern region

The attributes in the file used to compute surrogates include:

- oil – oil production in barrels
- gas – gas production in millions of cubic feet
- wells – number of wells
- spuds – number of spuds
- SHPID – filter for well type: A = all wells, O = oil wells, G = gas wells, C = CBM wells, S = spuds (drilling)

Shapefile Name: US_Oil_Gas_2008West2km_2011East4km

Type: Polygon

Year: 2008 and 2011

Attributes: oil, gas, wells, shp_id, spuds, SHPID

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699

2.7 Other Industrial and Commercial Activities

2.7.1 Mines

2003 USGS point locations of mineral operations and mines for the following: agricultural uses, construction uses, ferrous metal mines, ferrous metals processing plants, miscellaneous industrial uses, nonferrous metal mines, nonferrous metal processing plants, refractory, abrasive, and other industrial uses, sand and gravel, and crushed stone operations; NOTE: this Shapefile does not include energy minerals (i.e., coal is not included here).

The attributes in the file used to compute surrogates include:

- NONE – point locations of the mining operations are used to derive the surrogate

Shapefile Name: mines_usgs2003_wrf

Shapefile Type: Point

Year: 2003

Attributes: SNDGRVX020,COMP_NAME,OPER_NAME,COUNTY,STATE,FIPS,STATE_FIPS,LATITUDE,LONGITUDE,OPER_TYPE,COMMODITY

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 860, 861

2.7.2 Quarries

2003 USGS point locations of mineral operations and mines for the following: sand and gravel, and crushed stone operations

The attributes in the file used to compute surrogates include:

NONE – point locations of the mining operations are used to derive the surrogate

Shapefile Name: quarries_usgs2003_wrf

Shapefile Type: Point

Year: 2003

Attributes: SNDGRVX020,COMP_NAME,OPER_NAME,COUNTY,STATE,FIPS,STATE_FIPS,LATITUDE,LONGITUDE,OPER_TYPE,COMMODITY

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 862

2.7.3 Dry Cleaners

Counts of dry cleaners by Census Block Group in operation during calendar year 2009 were used to update the dry cleaner location surrogate. These data were purchased from the Spatial Insights, Inc. Business Counts Database¹⁰. As these data are proprietary, this Shapefile cannot be redistributed.

The attributes in the file used to compute surrogates include:

- DRY – number of dry cleaner operations in each Census Block Group

Shapefile Name: drycleaner_surrogate

Shapefile Type: Polygon

Year: 2009

Attributes: STFID, FIRST_STAT, FIRST_COUN, FIRST_TRAC, FIRST_BLOC, GEOID, DRY, FIPSSTCO

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 880

2.7.4 Gas Stations

Counts of gas stations (with and without convenience stores) by Census Block Group in operation during calendar year 2009 were used to update the gas station location surrogate. These data were purchased from the Spatial Insights, Inc. Business Counts Database⁹. As these data are proprietary, this Shapefile cannot be redistributed.

The attributes in the file used to compute surrogates include:

- GAS – number of gas station operations in each Census Block Group

Shapefile Name: gas_station_surrogate

Shapefile Type: Polygon

Year: 2009

¹⁰ <http://www.spatialinsights.com/catalog/product.aspx?product=126&content=1606>

Attributes: STFID, FIRST_STAT, FIRST_COUN, FIRST_TRAC, FIRST_BLOC, GEOID, GAS, FIPSSTCO

Projection: Lambert Conformal Conic ($X_0=-97.0$, $Y_1=33.0$, $Y_2=45.0$, $Y_0=40.0$)

Datum: NAD83

Spheroid: Normal Sphere with radius 6,370,000 meters

Derived Surrogates: 600

3 Known Issues and Future Work

While the Shapefiles and surrogates documented here and in the accompanying spreadsheet represent the best available information for allocating non-point emissions to modeling grids, there are some caveats to these data and issues that should be addressed in the future.

- Some of the surrogates are derived from proprietary Shapefiles that cannot be distributed. These data include:
 - The off-network mobile surrogates (major roads, bus stops, and highway exits) are generated from Shapefiles contained in the 2010 ESRI Data and Maps CD.
 - The dry cleaners and gas stations surrogates are generated from data purchased from Spatial Insights, Inc.
 - The oil and gas surrogates (689-699) are generated from data purchased from IHS, Inc. These data were aggregated from points to polygons (grid cells) to obscure the proprietary nature of the data. As described in Section 2.6, this aggregation process results in the loss of some of the original data.
- The off-network mobile surrogates (251-253) were generated using data that approximated the activity locations. As the off-network, heavy duty diesel truck rate per profile (RPP) emissions from MOVES simulates idling at truck stops and travel plazas, the surrogate for these sources should be derived from a point Shapefile that characterizes these locations. Although we were not able to find a national Shapefile of truck stops/travel plazas, these data are available for some states. An improvement of surrogate 251 (Off-Network MD/HD) will be to integrate a truck stop/travel plaza Shapefile into this surrogate.
- The total agriculture surrogate (310) is not specific enough for livestock sources and could be improved using a point Shapefile of confined animal feeding operations (CAFOs). Although we were not able to find a national Shapefile of CAFO locations, these data are available from some states. State departments of water quality track CAFO locations for monitoring runoff and water permitting purposes. It will be worthwhile to collect data from as many states as possible for developing a national CAFO surrogate to use for allocating non-point livestock inventories to model grids.
- The high resolution, national NCLD Shapefile breaks the Surrogate Tool by causing memory allocation failures while processing the data (SRGCREATE_ERROR>gpc malloc failure). To avoid this error it is necessary to split the NLCD Shapefiles into subregions, generate surrogates for each subregion, and then paste the data back together

to create a national surrogate. We needed to split the NLCD Shapefiles into eight subregions to avoid the memory violations in the Surrogate Tool. The Surrogate Tool needs to be updated to take advantage of the larger memory and multi-processing capabilities of modern computing hardware.

4 Surrogate Revision History

- Version 4 (09/2013) Updates
 - Shipping lanes surrogate (802): Built as a combination of Great Lakes shipping lanes/Inland waterways, Gulf of Mexico platform densities (803), and offshore shipping locations weighted by NEI2011 NOX emissions from SCC 2280002200 (806), and water (350).
 - Ports weighted by NEI2011 NOx emissions (820): Port areas joined with NEI2011 NOX emissions from SCC 2280002100
 - Oil and gas development surrogates for the following basins: Greater Green River (WY), Uintah (UT), Piceance (CO), Denver-Julesberg (CO), San Juan (NM), Raton (NM), Paradox (CO/NM), Permian (TX/NM), Marcellus (PA/NY/OH/WV/KY), Illinois (IL/IN); surrogates include Gas production at all wells (689), Oil production at all wells (690), Well count of CBM wells (691), Spud count (692), Well count of all wells (693), Oil production at oil wells (694), Well count of all oil wells (694), Gas production at gas wells (696), Oil production at gas wells (697), Well count of gas wells (698), Gas production at CBM wells (699).
 - Gas stations (600) - 2012 gas station point locations weighted by the number of locations in each Census Block Group
 - Dry cleaners (880) - 2012 dry cleaner point locations weighted by the number of locations in each Census Block Group
 - Great Lakes Tug Zone Area (808) - Tug zone from 0.5 - 2.0 miles offshore in the Great Lakes
- Version 3 (10/2012) Updates
 - All NLCD-based (3**) surrogates updated with NLCD 2006 30-m raster files aggregated to 250-m Shapefiles.
 - Includes new med. (301) and high intensity residential (302) surrogates
 - Includes new open space surrogate (303)
 - Total agriculture with orchards/vineyards and orchards/vineyards no longer available
 - Includes new pasture land (318) and crop land (319) surrogates
 - Four new housing surrogates (131-134) from the 2010 US Census data.
 - Three new rail density surrogates (271-273) from 2012 NTAD freight density line Shapefiles

- New rail yard surrogate (275) from ERTAC rail yard point Shapefile (provided by Georgia DNR)
- Construction and mining surrogate (861) added as a combination of 0.5 housing change/population + 0.5 mines
- Quarries surrogate (862) from point locations of mineral operations and mines for sand, gravel, and crushed stone
- Version 2 (06/2012) Updates
 - All FEMA-based (5**) surrogates re-generated with the correct data Shapefile (cty_pophu2k_revised). One artifact of using the incorrect data Shapefile was the inclusion of defunct counties/FIPs or the exclusion of new counties/FIPs in the new surrogates. Dade County, FL and Yellowstone National Park, MT were removed in the new data. Broomfield County, CO, Miami-Dade County, FL, and Boyd County, NE now exist.
 - All surrogates use new gapfilling definitions, including being gapfilled with surrogates that used the correct data Shapefile; some of the previous surrogates were gapfilled using FEMA data that used an incorrect data Shapefile
 - Two WY counties had surrogates that didn't normalize to 1.0. The Normalization Tool was used to force these two counties to have surrogates that add to 1.0.

5 REFERENCES

- Bergin, M. et al (2009) ERTAC Rail: A collaborative effort in building a railroad-related emission inventory between Eastern states' air protection agencies and participation with the railroad industry, presented at the 18th Annual Emission Inventory Conference, U.S. EPA, Baltimore, MD, <http://www.epa.gov/ttnchie1/conference/ei18/session6/bergin.pdf>.
- Eyth, A. et al. (2007) Accounting for Land Use Changes in Projecting Future-Year Emissions Scenarios, presented at the 16th Annual Emission Inventory Conference, U.S. EPA, Raleigh, NC, <http://www.epa.gov/ttnchie1/conference/ei16/session7/eyth.pdf>.