www.erg.com

**TECHNICAL MEMORANDUM**

**TO**: Alison Eyth and Laurie Trinca
Delivery Order Managers
U.S. EPA Office of Air Quality Planning and Standards

**CC**: Ted Palma, Jennifer Snyder, Madeleine Strum, and Jeffrey Vukovich
U.S. EPA Office of Air Quality Planning and Standards

**FROM**: Regi Oommen, Rebecca Bayham, Lindsay Dayton, Bebhinn Do, and Heather Perez, Eastern Research Group, Inc. (ERG)

**DATE**: September 2, 2016

**SUBJECT**: Modeling Allocation Factors for the 2014 Oil and Gas Nonpoint Tool

---

## 1.0 INTRODUCTION

The exploration and production of oil and gas has increased in terms of quantities and locations over the last five years, primarily through the use of new technologies, such as hydraulic fracturing. As part of the 2011 National Emissions Inventory (NEI) cycle, EPA recently prepared county-level emission estimates for the oil and gas sector. This emissions inventory was more comprehensive on a geographic, source category, and pollutant coverage basis when compared to prior NEI base years for this sector. Under EPA Contract No. EP-D-09-048, Delivery Order 00-54, entitled "Report Development – Data Characterization", ERG prepared census-tract and 4-km sub-county surrogate factors for 15 surrogates for EPA to use in emissions modeling.

The purpose of this memorandum is to summarize procedures used to develop spatial and temporal modeling allocation factors for the 2014 Oil and Gas Nonpoint Tool using data primarily from a third-party database of oil and gas wells, and other sources. EPA directed ERG to start with the analysis and files delivered to EPA Office of Compliance on August 1, 2015 (U.S. EPA, 2015), and to incorporate additional datasets to develop surrogate modeling factors. All work was performed under EPA Contract No. EP-D-14-030, Delivery Order 00-27, entitled "Data Analysis/Report Development."

## 2.0 BACKGROUND INFORMATION

EPA uses the NEI for several purposes, including emissions modeling for regulatory activities. In support of 2014 baseyear emissions for the NEI, EPA recently developed Version 1 of the 2014 National Oil and Gas Emissions Estimation Tool (U.S. EPA, 2016b). Although the activity data inputs in the Tool were at the county-level, much of the data originated from monthly well-level data that can be used for sub-county spatial and monthly temporal modeling. Additionally, through the development of the Tool, states had the opportunity to revise county-level activity data. For example, the Texas Commission on Environmental Quality (TCEQ) provided revisions to the original oil and gas well counts in the Tool. For other states, such as Kentucky, 2014 production data were only available at the state level, but were allocated to the county-level based on well-level data and surrogates from Kentucky's Oil and Gas Commission. EPA prepared documentation on the development of the 2014 activity data within the Tool (U.S. EPA, 2016b).

For this Delivery Order, ERG developed spatial allocation factors at both the 2-km and the 4-km grid scale level. Further, ERG was able to separate natural gas production activities from coalbed methane (CBM) activities to better match the level of specificity from EPA's Oil and Gas Tool. Additionally, ERG developed 2-km and 4-km shapefiles for Alaska, which did not have spatial allocation factors at these levels for the state. Finally, EPA asked ERG to develop monthly temporal allocation factors by SCC, which can be useful for future air quality modeling.

## 3.0 DATA SOURCES

The modeling surrogates were developed using multiple data sources described below.

### 3.1 HPDI

The primary activity data source used for the development of the oil and gas spatial surrogates was data from Drilling Info (DI) Desktop's HPDI database (Drilling Info, 2015). This database contains well-level location, production, and exploration statistics at the monthly level. Due to a proprietary agreement with DI Desktop, individual well locations and ancillary production cannot be made publicly available, but aggregated statistics are allowed. For the Tool, the individual well-level statistics were summed to the county-level. HPDI data represents nearly 94% of the activity data used in the Tool.

### 3.2 Oil and Gas Commission Websites

For the remaining 6%, ERG supplemented the HPDI activity data with additional data from Oil and Gas Commission (OGC) websites. In many cases, the correct surrogate parameter was not available (e.g., feet drilled), but an alternative surrogate parameter was available (e.g., number of spudded wells) and downloaded. The types of information retrieved from these websites are presented in Table 1, as well as the corresponding reference listed in Section 8.

**Table 1. Information Retrieved from State Websites**

| State | Information Retrieved | Reference |
|---|---|---|
| Arizona | Well Locations, Spud Counts, Well Depths | Arizona OGC, 2015 |
| Idaho | Well Locations, Spud Counts, Well Depths, Gas Production, Produced Water, Well Completions | Idaho OGC, 2015 |
| Illinois | Well Locations, Spud Counts, Well Depths, Oil Production, Gas Production, Well Completions | Illinois SGS, 2016 |
| Indiana | Well Locations, Spud Counts, Well Depths, Oil Production, Gas Production, Well Completions | Indiana OGC, 2016 |
| Kentucky | Well Locations, Spud Counts, Well Depths, Oil Production, Gas Production, Produced Water, Well Completions | Kentucky GS, 2016 |
| Missouri | Well Locations, Spud Counts, Well Depths, Oil Production, Gas Production, Produced Water | Missouri DNR, 2015 |
| Nevada | Well Locations, Spud Counts, Well Depths, Well Completions | Nevada DMR, 2015 |
| Oregon | Well Locations, Gas Production | Oregon OGC, 2016 |
| Pennsylvania | Well Locations, Produced Water | Pennsylvania DEP, 2015 |
| Tennessee | Well Locations, Spud Counts, Well Depths, Gas Production, Well Completions | Tennessee DEP, 2015 |

### 3.3 EPA Greenhouse Gas Inventory for Completions

EPA supplemented the completion information from HPDI by implementing the methodology for counting oil and gas well completions developed for the U.S. National Greenhouse Gas Inventory (U.S. EPA, 2013). Under that methodology, both completion date and date of first production from HPDI were used to identify wells completed during 2014.

## 4.0 DATA COMPILATION

In total, over 1.43 million unique wells were compiled from the above data sources. The wells cover 34 states and 1,158 counties. Well locations are presented in Figure 1. Each well was

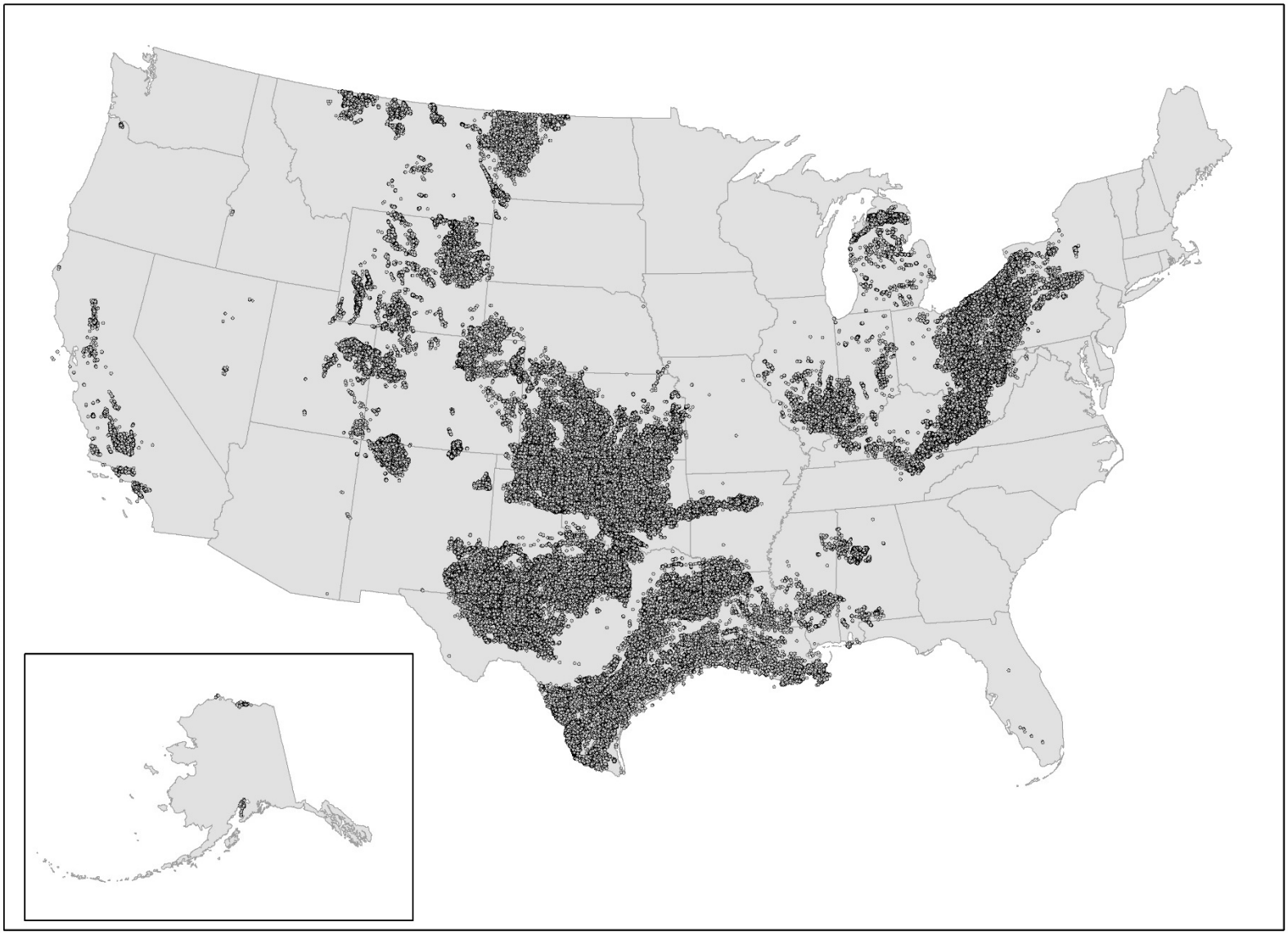

**Figure 1. Compiled Well Locations for the U.S**

uploaded into ArcGIS, and assigned to the associated 2-km and 4-km grid identifier. The 4-km grid was provided by EPA (EPA, 2016a), and ERG developed a 2-km grid using the 4-km grid.

For the development of sub-county modeling surrogates, attribute data (e.g., production, well counts, produced water, etc.) for each well were assigned both 4-km and 2-km modeling grid identifiers. By default, attribute data were initially summed to the 2-km modeling grid level. If the data for the attribute was based on less than 3 wells within the 2-km modeling grid, then the wells were summed to the 4-km modeling grid. For the majority of the attributes, wells remained in the 2-km modeling bins. Table 2 summarizes the well counts of the 23 attributes by modeling grid.

**Table 2. Oil and Gas Surrogate Codes**

| Oil and Gas Attribute | Number of Wells in 2-km Modeling Grid | Number of Wells in 4-km Modeling Grid | Total Number of Wells |
|---|---|---|---|
| Associated Gas Production | 465,326 | 39,435 | 504,761 |
| CBM Production | 42,571 | 2,962 | 45,533 |
| CBM Well Counts | 42,571 | 2,962 | 45,533 |
| Completions – All Wells | 26,167 | 13,595 | 39,762 |
| Completions – CBM Wells | 122 | 189 | 311 |
| Completions – Gas Wells | 3,270 | 2,471 | 5,741 |
| Completions – Oil Wells | 22,308 | 11,402 | 33,710 |
| Condensate Production – CBM Wells | 870 | 532 | 1,402 |
| Condensate Production – Gas Wells | 62,894 | 18,023 | 80,917 |
| Feet Drilled | 23,363 | 17,049 | 40,412 |
| Gas Production | 365,525 | 51,552 | 417,077 |
| Gas Well Counts | 365,525 | 51,552 | 417,077 |
| Oil Production | 874,098 | 56,696 | 930,794 |
| Oil Well Counts | 874,098 | 56,696 | 930,794 |
| Produced Water – All Wells | 820,324 | 43,032 | 867,356 |
| Spud Counts – All Wells | 26,877 | 1,353 | 28,230 |
| Spud Counts – CBM Wells | 122 | 193 | 315 |
| Spud Counts – Gas Wells | 3,262 | 2,442 | 5,704 |
| Spud Counts – Oil Wells | 23,030 | 11,363 | 34,393 |
| Total Exploratory Wells | 27,366 | 13,945 | 41,311 |
| Total Production Wells | 1,315,657 | 77,747 | 1,393,404 |
| Total Wells | 1,353,335 | 77,407 | 1,430,742 |
| Unconventional Well Completions | 17,245 | 8,260 | 25,505 |

Figure 2 presents the combined 2-km and 4-km modeling grid coverages.

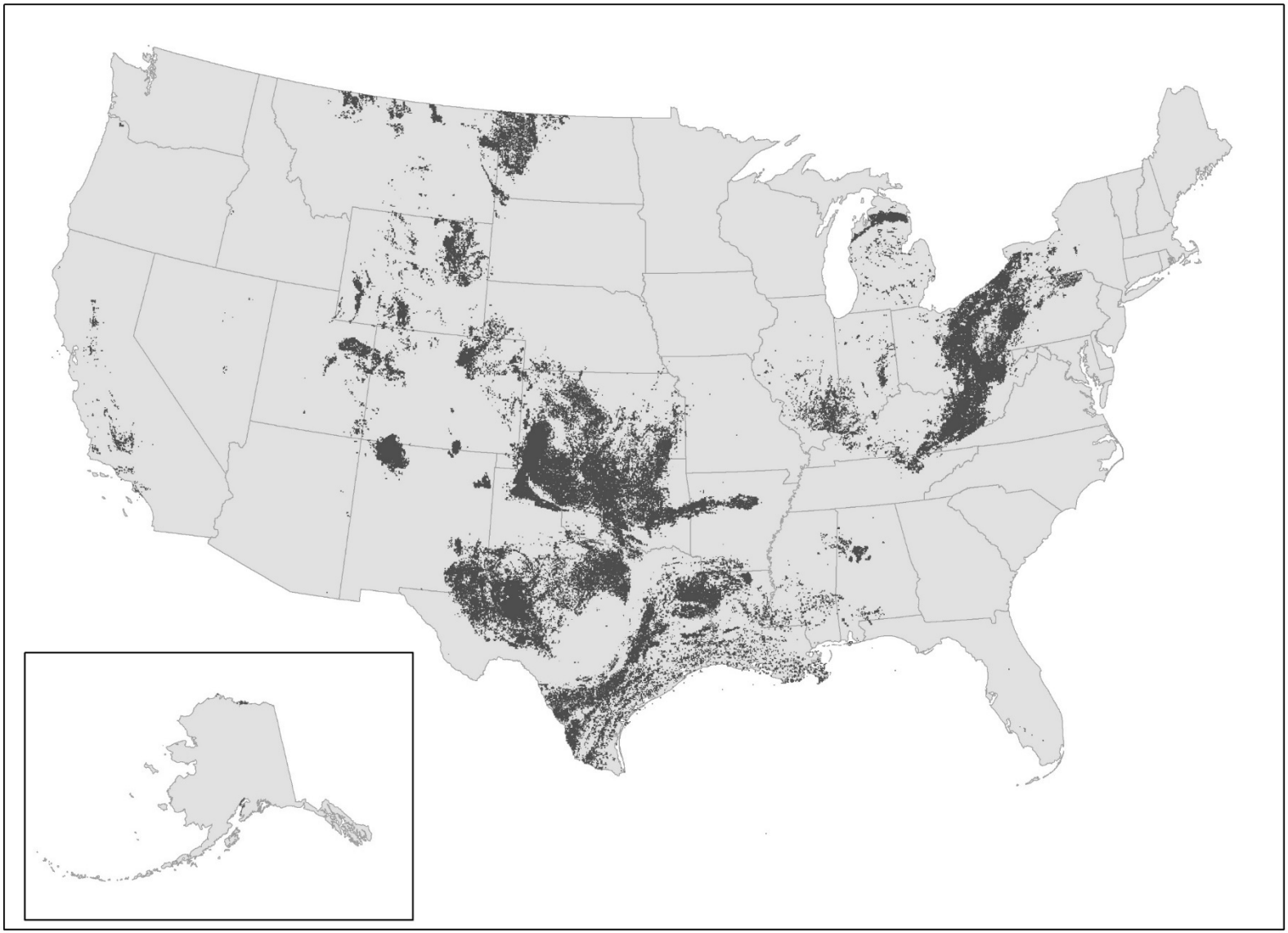

**Figure 2. Compiled Well Locations Placed at U.S. 2-km and 4-km Grids**

## 5.0 OIL AND GAS SURROGATES

The 2014 Nonpoint Oil and Gas Emissions Estimation Tool contains emission estimates for 34 states and 1,158 counties. Additionally, emissions are estimated for 54 oil and gas source classification codes (SCCs), those that begin with 2310xxxxxx. The list of SCCs from the Tool is presented in Appendix A. In total, there are 29,162 unique county-SCC pairs with emissions.

Despite the large number of SCCs, emission surrogates were allocated to the 2-km and 4-km level for twenty-three surrogates. These surrogates are presented in Table 3.

**Table 3. Oil and Gas Surrogate Codes**

| EPA Surrogate Code | New Surrogate for 2014 NEI | EPA Surrogate Description | Surrogate Spatial Allocation Factor Name |
|---|---|---|---|
| 670 | YES | Spud count - CBM Wells | SPUD_CBM |
| 671 | YES | Spud count - Gas Wells | SPUD_GAS |
| 672 | YES | Gas production at Oil wells | ASSOCIATED_GAS_PRODUCTION |
| 673 | YES | Oil production at CBM Wells | CONDENSATE_CBM_PROD |
| 674 | YES | Unconventional Well Completion Counts | SPUD_HF |
| 676 | YES | Well count - all producing | TOTAL_PROD_WELL |
| 677 | YES | Well count - all exploratory | TOTAL_EXPL_WELL |
| 678 | YES | Completions at Gas Wells | COMPLETIONS_GAS |
| 679 | YES | Completions at CBM Wells | COMPLETIONS_CBM |
| 681 | | Spud count - Oil Wells | SPUD_OIL |
| 683 | | Produced Water at all wells | PRODUCED_WATER_ALL |
| 685 | | Completions at Oil Wells | COMPLETIONS_OIL |
| 686 | | Completions at all wells | COMPLETIONS_ALL |
| 687 | | Feet drilled at all wells | FEET_DRILLED |
| 691 | YES | Well counts - CBM Wells | CBM_WELL |
| 692 | | Spud count - All Wells | SPUD_ALL |
| 693 | | Well count - all wells | TOTAL_WELL |
| 694 | | Oil production at oil wells | OIL_PRODUCTION |
| 695 | | Well count - oil wells | OIL_WELL |
| 696 | YES | Gas production at Gas wells | GAS_PRODUCTION |
| 697 | | Oil production at Gas Wells | CONDENSATE_GAS_PROD |
| 698 | | Well counts - Gas Wells | GAS_WELL |
| 699 | YES | Gas production at CBM wells | CBM_PRODUCTION |

.

Due to a change in methodology for the 2014 NEI, hydraulic fracturing emissions are now based on unconventional well completion counts. For the 2011 NEI, hydraulic fracturing emissions were based on count of horizontally-spudded wells, and so this surrogate was not developed for the 2014 NEI. Additionally, as a result of better specificity separating activities of gas wells from CBM wells, three combined surrogate codes were not developed for the 2014 NEI. These surrogates are presented in Table 4.

**Table 4. 2011 Oil and Gas Surrogate Codes Not Developed for the 2014 NEI**

| EPA Surrogate Code | EPA Surrogate Description | Reason for Not Developing |
|---|---|---|
| 682 | Spud count – Horizontally-drilled wells | New allocation code developed for hydraulically-fractured wells based on Unconventional Well Completions (Surrogate Code = 674) |
| 684 | Completions at Gas and CBM Wells | Developed separate spatial allocation factors for completions at gas wells (Surrogate Code = 678) and completions at CBM wells (Surrogate Code = 679) |
| 688 | Spud Count – Gas and CBM Wells | Developed separate spatial allocation factors for spud counts at gas wells (Surrogate Code = 671) and spud counts at CBM wells (Surrogate Code = 670) |
| 689 | Gas Production at All Wells | Developed separate spatial allocation factors for gas production at gas wells (Surrogate Code = 696), gas production at oil wells (Surrogate Code = 672), and gas production at CBM wells (Surrogate Code = 699) |

Finally, Table 5 presents 2011 oil gas surrogate codes that changed descriptions as a result of the increased specificity of data for the 2014 NEI.

**Table 5. Updates Surrogate Code Descriptions for the 2014 NEI**

| EPA Surrogate Code | 2011 Surrogate Description | 2014 Surrogate Description |
|---|---|---|
| 697 | Oil Production at Gas and CBM Wells | Oil Production at Gas Wells |
| 698 | Well Counts – Gas and CBM Wells | Well Counts – Gas Wells |

Appendix B presents the county-SCC pairs with the primary surrogate codes. If the primary surrogate was not available, then an alternate surrogate was assigned. Appendix C presents the surrogate assignment progression. In cases where there is no well-level location data

for a particular county, then a surrogate code of 400, which is allocation by rural land area, was assigned as an alternative surrogate.

## 6.0 SPATIAL SURROGATE CALCULATIONS

Since nonpoint oil and gas emissions are at the county-level, the surrogate factors need to be developed for portions within the county.

### 6.1 Spatial Surrogate Calculations – 4-km Grid Scale

For the 4-km allocation, the following steps were used:

a. Using GIS software, assign wells to 4-km grid cell
b. Sum allocation factors to the county- and 4-km grid cell-level
c. Sum allocation factors to the county-level
d. Divide summed county- and 4-km grid cell allocations by the summed county-level allocations to calculate 4-km spatial allocation factors

For example, County B reports 25 tons of CO from 4-stroke rich burn Wellhead Compressors from 1,000 natural gas wells. In an adjacent county, County C reports 15 tons of CO from 100 natural gas wells for the same source category. The two counties share one similar 4-km grid cell (55) on the border. The primary activity data used to generate emissions is the number of natural gas wells. County B has a total of 100 natural gas wells in grid cell 55 and County C a total of 50 natural gas wells in grid cell 55.

Thus, CO emissions for grid cell 55 within County B are then apportioned, as follows:

**Step 1**

$$\text{County B, grid cell 55 CO emissions} = \frac{(25\ \text{tpy, County A}) * (100\ \text{natural gas wells, Count B, grid cell 55})}{1{,}000\ \text{natural gas wells in County B}}$$

County B, grid cell 55 CO emissions = 2.5 tpy

**Step 2**

$$\text{County C, grid cell 55 CO emissions} = \frac{(15\ \text{tpy, County A}) * (50\ \text{natural gas wells, Count C, grid cell 55})}{100\ \text{natural gas wells in County B}}$$

County C, grid cell 55 CO emissions = 7.5 tpy

**Step 3**

Grid cell 55 CO emissions=2.5 tpy+7.5 tpy=10.0 tpy

Over 445,000 Surrogate Code-County-Grid Cell IDs were developed, and are presented in Appendix D.

## 6.2 Temporal Surrogate Calculations

Monthly surrogates were prepared for county-SCC combinations which overlap with data extracted from HPDI, state Oil and Gas Commission websites, and RIGDATA. The following steps were used to generate the monthly surrogates:

a. Sum allocation factors to the monthly timeframe
b. Sum allocation factors to annual timeframe
c. Divide summed monthly allocations by the summed annual allocations to calculate monthly spatial allocation factors

For county-SCC combinations that were not extracted from HPDI, the surrogate parameter was evenly distributed by month. Appendix E presents the temporal factors by the county-SCC combinations.

## 7.0 FINAL DATA PRODUCTS

Final data products for this effort include:

- Appendix E – Monthly temporal factors by County and SCC. The temporal allocation factors were in one-record per line (ORL) format with the following data fields, and are presented in Appendix E:
    - FIPS
    - SCC
    - JANFRAC
    - FEBFRAC
    - MARFRAC
    - APRFRAC
    - MAYFRAC
    - JUNFRAC
    - JULFRAC
    - AUGFRAC
    - SEPFRAC
    - OCTFRAC
    - NOVFRAC
    - DECFRAC

- Appendix F – Shapefiles of the Well Attributes for the Continental U.S. (CONUS) and Alaska
- Appendix G – 4-km surrogate code text files for the CONUS only. For the 4-km spatial allocation factors, ERG prepared SMOKE-ready files for 23 surrogate codes. The 4-km spatial allocation factor files contain the following data fields:
    - Header Descriptions
    - Surrogate Code
    - State and County FIPS Code
    - Grid-Scale Column Value
    - Grid-Scale Row Value
    - Spatial Allocation Factor
    - Fractionated grid-level total value
    - County-level total

- Appendix H – CONUS and Alaska 2-km and 4-km merged county-level heat maps for 23 well attributes.

## 8.0 REFERENCES


Arizona Oil and Gas Commission (OGC). Recent Permits. Accessed 11/9/2015. Internet address: http://www.azogcc.az.gov/permits

Idaho Oil and Gas Commission (OGC). Idaho Oil and Gas Commission, Dept. of Lands. Activity Map. Accessed 11/9/2015. Internet address: http://www.idl.idaho.gov/oil-gas/leasing/index.html

Drillinginfo, Inc. 2015. "DI Desktop Database powered by HPDI." Internet address: http://www.didesktop.com/

Illinois State Geological Survey (SGS). Location Points from the ISGS Wells and Borings Database. Accessed 6/2/2016. Internet address: https://clearinghouse.isgs.illinois.edu/data/geology/location-points-isgs-wells-and-borings-database

Indiana Oil and Gas Commission (OGC). Oil & Gas Online Well Records Database. Accessed 6/8/2016. Internet address: http://www.in.gov/dnr/dnroil/5447.htm

Kentucky Geological Survey (GS). Search the Oil and Gas Records Database. Accessed 6/2/2016. Internet address: hhttp://kgs.uky.edu/kgsweb/DataSearching/OilGas/OGSearch.asp

Missouri Department of Natural Resources (DNR). Data and Reports, Oil and Gas Production. Accessed 11/10/2015. Internet address: http://dnr.mo.gov/geology/geosrv/oilandgas.htm

Nevada Division of Mineral Resources (DMR). Oil and Gas Permits. Accessed 11/9/2015. Internet address: http://minerals.nv.gov/Programs/Oil_and_Gas/OandG_Permits/Permits/

Oregon Oil and Gas Commission (OGC). Oil & Gas Well Log Viewer. Accessed 6/7/2016. Internet address: http://www.oregongeology.org/mlrr/spreadsheets/OG_Permits_02-10-2016.xls

Pennsylvania Department of Environmental Protection (DEP). Oil & Gas Reporting Website, Waste Reports by County. Accessed 11/6/2015. Internet address: https://www.paoilandgasreporting.state.pa.us/publicreports/Modules/Waste/WasteByCounty.aspx

Tennessee Department of Conservation. Oil and Gas Well Permits. Accessed 10/6/2015. Internet address: http://tdec.tn.gov:8080/pls/enf_reports/f?p=9034:34300:0:NO

U.S. EPA, 2013. Oil Well Counts and Completion Counts in Conventional and Unconventional Formations. Memorandum Prepared by Eastern Research Group, Inc. to Ms. Suzie Waltzer. 5/24/2013.

U.S. EPA, 2015. Land-Based Oil and Gas Extraction Knowledge Base Underlying Data from DrillingInfo, Inc. Prepared by Eastern Research Group, Inc. to Mr. Carey Johnston, U.S. EPA Office of Compliance. 10/15/2015.

U.S. EPA, 2016a. 4-KM Modeling Grid Description. Grid Definitions sent from Ms. Alison Eyth, U.S. EPA to Mr. Regi Oommen, Eastern Research Group via e-mail. 5/26/2016.

U.S. EPA, 2016b. Estimating Nonpoint Emissions from the Oil and Gas Production Sector, Revised Draft. Prepared by Eastern Research Group, Inc. 7/28/2016.

# Appendix A – 2014 Nonpoint Tool Oil and Gas SCCs

(See APPENDIX_A_2014_NONPOINT_SCC.xlsx)

# Appendix B – 2014 Nonpoint Oil and Gas Tool County FIPS, SCCs, and Surrogate

(see APPENDIX_B_2014_TOOL_COUNTY_SCC_SURROGATE.xlsx)

## Appendix C – Surrogate Priority

(See APPENDIX_C_SURROGATE_PRIORITY.pdf)

## Appendix D – Surrogate, County FIPs, and 4-km Grid Cells

(see APPENDIX_D_SURROGATE_COUNTY_GRID.xlsx)

## Appendix E – Monthly Temporal Factors

(see APPENDIX_E_MONTHLY_TEMPORAL_FACTORS.xlsx)

## Appendix F – CONUS and Alaska GIS Shapefiles

(see APPENDIX_F_CONUS_AK_ATTRIBUTE_SHAPEFILES.zip)

## Appendix G - 4-km Surrogate Modeling Files

(see APPENDIX_G_4_KM_SURROGATE_FILES.zip)

# Appendix H – Well Attribute Heat Maps

(see APPENDIX_H_WELL_ATTRIBUTE_HEAT_MAPS.zip)