Appendix 4.A

Methods to Determine Numbers and Locations of Hydraulically Fractured Wells in the Los Angeles Basin

4.A.1. Study Area

The geographic focus of this case is limited to the California Air Resources Board (CARB) South Coast Air Basin (SoCAB). Using GIS techniques, all datasets were limited to the boundaries of the SoCAB, which includes Los Angeles County, Orange County, and parts of both Riverside and San Bernardino counties (State of California, 2014).

4.A.2. Dataset Development

To create the datasets for the Los Angeles Basin (SoCAB) case study, we started with the analysis of well stimulation treatments in Volume I of this report. The data were refined so that all wells included in the analysis were, according to the Division of Oil, Gas, and Geothermal Resources (DOGGR), active oil and/or gas production wells that were stimulated using matrix acidizing, hydraulic fracturing, and frac-packing techniques. The Volume I dataset was therefore cross-referenced with the following datasets outlined in Table 4.A-1:

Table 4.A-1. Data Sources

Data Source #	Data Source	Data Location	Database Access Date
1	DOGGR All Wells database	http://www.conservation.ca.gov/dog/maps/Pages/GISMapping2.aspx	12/14/14
2	DOGGR SB4 Well Simulation Notices	http://maps.conservation.ca.gov/doggr/iwst_index.html	12/14/14
3	DOGGR Well Production Database	http://www.conservation.ca.gov/dog/prod_injection_db/Pages/Index.aspx	12/14/14
4	SCAQMD Rule 1148.2 Oil and Gas Well Electronic Notification and Reporting	http://www.aqmd.gov/home/regulations/compliance/1148-2	12/14/14
5	FracFocus 1 and FracFocus 2	See Approach in Volume I	

Sources 1-3 were used to identify which wells we consider to be "Producing" (oil/gas) and the producing wells that have been "Stimulated." Errors in the data and missing information complicated the process, but the methodology used in this case study was as follows:

1. Active Oil and/or Gas Production Wells Dataset: The DOGGR production database was joined to the DOGGR AllWells dataset using a common field, the American Petroleum Institute (API) identifier.
    a. The data was sorted by most recent year of production.
    b. All wells with production numbers from 2013/2014 were included.
2. Wells that had not reported production (oil or gas) figures for the last five years were sorted by multiple fields for consideration.
    a. Wells were first sorted by "GIS Symbol Codes." (see Table 4.A-1) Wells with codes identifying the well as an "active" or "new" oil and/or gas wells were included. These were wells marked as AOG (Active Oil and Gas), ADG (Active Dry Gas), NOG (New Oil and Gas) and NDG (New Dry Gas). "New" wells are somewhere in development process between being permitted and inspected/approved for production.

Table 4.A-2. Three letter DOGGR code representing the well status (first letter in code), and the well type (last two letters in code).

1st Letter	Definition
N	New (permitted to drill)
A	Active or Idle
P	Plugged

+

2nd and 3rd Letter	Definition
AI	Air Injector
DG	Dry Gas
GD	Gas Disposal
GS	Gas Storage
LG	Liquid Petroleum Gas
OB	Observation
OG	Oil & Gas
PM	Pressure Maintenance
SF	Steam Flood
WD	Water Disposal
WF	Water Flood
WS	Water Source
NK	Unknown
DH	Dry Hole

Wells were then sorted by the "Well Status Code" (see Table 4.A-2). Wells that reported production for 2014 were included in our "Production Wells" dataset. DOGGR codes for well status are described in Table 4.A-3.

Table 4.A-3. Definition of well status per DOGGR

Well Status	Definition	Explanation
N	New	Recently permitted, in the process of being drilled.
B	Buried	Older Well, not abandoned to today's standards, location of well is approximate.
U	Unknown	Status not yet entered from hard copy file. Wells are mostly older, pre-1976.
A	Active	Well has been drilled and completed
C	Cancelled	Well permit was cancelled -- well not drilled
P	Plugged and Abandoned	Well gas been plugged and abandoned to Division standards
I	Idle	Well is idle, not producing

a. Wells without reported production figures for 2013/2014 and marked as B, U, C, P, or I were excluded from further analysis resulting in the following total set of active production wells:

i. California: 74,482 wells

ii. South Coast Air Basin: 4,487

iii. Los Angeles County: 4,068

4.A.3. Stimulation Dataset Development

1. The SB4 Analysis Volume I dataset was cleaned and edited, leaving a total of 4,780 unique entries with lat/lon coordinates for the state of California.

   a. SB4 well stimulation completion records were added to the dataset.

   b. An additional 8 wells from FracFocus were added to the stimulations dataset. They were:

FracFocus	0403053941	35.614947	-119.721269
FracFocus	0403054276	35.436805	-119.691225
FracFocus	0403054279	35.440473	-119.693172
FracFocus	0403054280	35.437066	-119.689655
FracFocus	0403054281	35.435165	-119.694609
FracFocus	0403054304	35.461401	-119.727668
FracFocus	0403053942	35.615100	-119.720973
FracFocus	0403053943	35.614878	-119.721955

2. For the Los Angeles Basin analysis, known stimulated wells were limited to the boundaries of the South Coast Air Basin (SoCAB). Of the total (n1=199),

139 were identified as hydraulically fractured and 60 were identified as matrix acidizing.

3. The Volume I analysis developed stimulation probabilities for individual production pools.

    a. For pools with stimulation probabilities >50%, it was assumed that all wells in the pools were stimulated. These wells were added to the known stimulations dataset, for a total of (n2=1256)

4. For pools with stimulation probabilities <50% (n3=1462), the probabilities were applied to the number of active production wells in the pool, for a total (N= 1289):

    a. Las Cienegas pool (33%) 28 active production wells

    b. Wilmington Lower Terminal (33%) 1 active production well

    c. Wilmington (17%) 47 active production wells

    d. Tapia (7%) 30 active production wells

    e. Wilmington Ranger (6%) 55 active production wells

    f. Wilmington Tar (2%) 45 active production wells

Comparing across methods had a trivial impact on estimated total number of individuals and attributes of the population within specified buffer distance. At all buffer distances considered, Groups 2 and 3 included almost identical human population numbers, which were approximately 15% greater than Group 1 (Figure 4.A-1). For example, the total population estimate within the 2,000 m (6,562 feet) buffer distance was 668,631 for Group 1; 752,810 for Group 2; and 759,512 for Group 3. Considering the similar population estimates, and to be inclusive of all individuals likely to be exposed at fine spatial scales, analyses of demographic characteristics focused on the Group 3 estimation method.

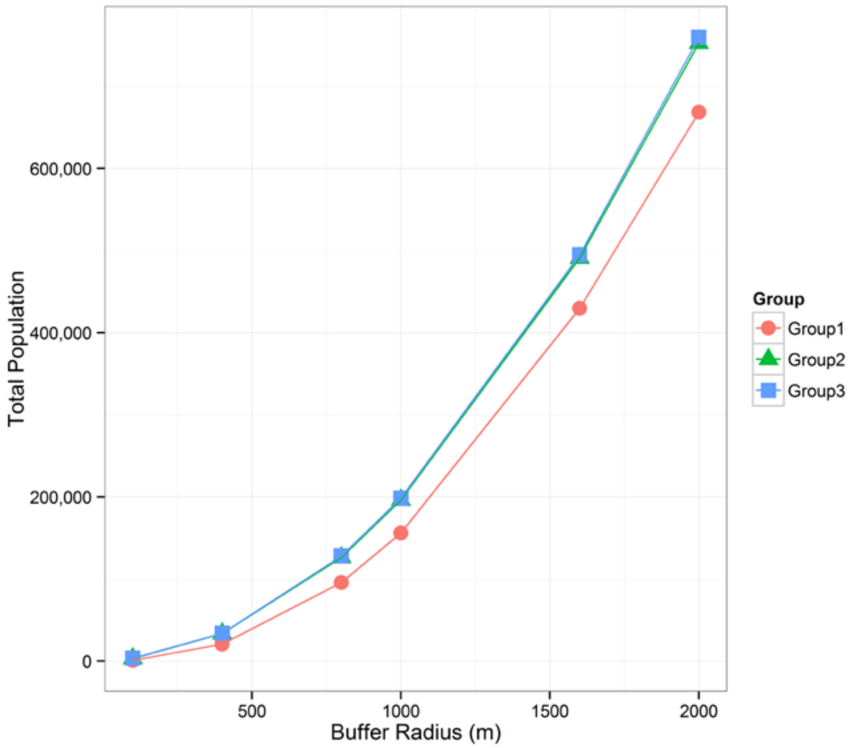

Figure 4.A-1. Total human population numbers estimated within each distance radius, employing each of the three considered grouping methods. Group 1=all known and confirmed wells that have been stimulated; Group 2=(Group 1) + (all probabilities of pools with >50% stimulation records); Group 3=(Group 2) + (all pools with <50% of wells reported as stimulated with applied adjustment).

Appendix 4.B

Methods and Approach for Spatial Analysis of Potential Public Health Risks

4.B.1. Data Sets Used

4.B.1.1. U.S. Census Bureau Data

Decennial census data (U.S. Census Bureau, 2010) was downloaded from American Fact Finder via Census.gov for the entire state of California at the census block and the American Community Survey (ACS) (2013 five-year estimates) was downloaded at the block group level (U.S. Census Bureau, 2010: 2012; U.S. Census Bureau).

4.B.2. Sensitive Receptors

In addition to demographic profiles of residents, we included four classifications of sensitive receptor points to our analysis. These locations represent sites where a hazard may present an elevated risk to populations that are known to be more vulnerable than the general population with regards to air quality and environmental degradation exposures. For the South Coast Air Basin (SoCAB) study area, locations of residential elderly care homes, elementary and secondary schools, permitted daycare facilities, and playgrounds were included.

Data sets included in the sensitive receptors analysis include:

4.B.3. Residential Elderly Care Homes

This dataset was taken from the California Health Care Facility Dataset (HLTHFAC); a dataset of over 4,000 facilities in California. The dataset was limited to only residential elderly care facilities (State of California Geoportal, 2014).

4.B.4. Schools

Enrollment demographics data was downloaded from the California Department of Education website. The dataset of schools from the CA.gov Geoportal was downloaded, cleaned and the locations verified for elementary, secondary, and unified school districts. Then 2013/2014 enrollment demographic profiles of each school was joined to the shapefile (U.S. Census Bureau, 2014). Schools with (0) total student enrollment were removed from the dataset. Quality control techniques identified enrollment demographics for schools that did not match the schools listed in the GIS files, which were eliminated from the analysis (California Department of Education, 2014; U.S. Census Bureau, 2014).

4.B.5. Permitted Daycare Facilities

This dataset consists of licensed childcare centers/preschools or daycares. It was extracted from the larger dataset of all childcare facilities, which also includes child group homes (California Department of Social Services, 2014). The dataset of playgrounds was limited from a dataset of all recreational facilities in California's cities/urban centers. The dataset was provided by the San Francisco-based non-profit, GreenInfo Network - Information and Mapping in the Public Interest (GreenInfo Network, 2012).

4.B.6. Spatial Analysis Methods

All datasets were clipped to the study area of the SoCAB using GIS techniques in ArcView 10.2.2 software. The datasets were all projected in the NAD 1983 California Teale Albers coordinate system. Projected shapefiles were then imported into a personal geodatabase for ease of management and sharing with other Arc users.

Determination of a "stimulated, active oil and gas production well" dataset for an assessment of vulnerable populations required an adherence to inclusivity. To create the dataset, all pools with known stimulation events were included, and the probability of well stimulations within the pools and within the surrounding communities. If pools estimated to have <50% wells stimulated were excluded, communities and important demographics would also be excluded from the analysis. Active oil and gas production wells located within pools <50% stimulated were therefore considered in the analysis. The well stimulation probabilities were applied as adjustment factors to the demographics of the surrounding communities and to the number of wells from said pool falling within the boundaries of sensitive receptors. If an area/community fell within the bounds of multiple buffers of "pools <50%", the stimulation probabilities were summed to generate an appropriate adjustment values.

Socio-economic characteristics of communities living at various proximities to active and stimulated active production wells were analyzed and summarized using a combination of GIS and spreadsheet manipulation techniques. Select demographics for census blocks were extracted from the 2010 decennial U.S. census data, and additional demographical data was extracted from the 2010 decennial census at the block group level. The demographics selected for the analysis available at the census block resolution included total population counts, racial demographics, and age profiles. From the counts by race, "non-Hispanic minority", "Hispanic population" and "total minority population" fields were calculated. From the age profiles, counts of individuals "5 years of age and younger", "<18 years of age" and "75 years of age and older" were calculated. For data relating to economic status, the highest resolution collected by the U.S. census bureau is limited to census block groups and is collected during the ACS. Economic status data by block group was joined to the census block level data using a common field in both datasets, the GEOID for census block groups. The ACS (2013, 5-year) block group data included household counts, educational attainment, limited English language households, and household income, poverty status and employment. From this data the following indicators of socio-economic status were generated: “no high school graduation”, “limited English speaking household”, “median income”, “income below poverty line”, “income below individual household food-stamp requirement”, and “unemployed”.

For the analysis of vulnerable population characteristics, the dataset of block and block group census data was sampled at multiple distances from all active wells and stimulated active wells. The dataset of “stimulated, active oil and gas production wells” was buffered at multiple radial distances: 100 m (328 feet), 400 m (1,312 feet), 800 m (2,625 feet), 1,000 m (3,281 feet), 1,600 m (5,249 feet), and 2,000 m (6,562 feet) radii, as discussed in this chapter and in Volume II, Chapter 6.

The census data was mapped using GIS techniques; the census data was joined to a map layer of census blocks within the bounds of the SoCAB. The census blocks were then clipped to the bounds of each buffer radius. The areal extent of these clips was calculated for each census block. For blocks that fell across the boundary of a buffer, the total area of the block divided by the area of the block falling within the buffer was calculated, and this areal percent was applied to the census block demographic data. For the block group resolution census counts, an areal percentage equal to the clipped area of the census block divided by the total block group area was applied to demographics data. For the census blocks and block groups falling within the buffer of a pool with <50% of the wells stimulated, the well stimulation probability of that pool was then applied to the census counts. The population estimates of the sample population were then compared to the demographics for the rest of the SoCAB area, which was defined as the population living greater than 2,000 m (6,562 feet) from stimulated wells.

An analysis of four classes of sensitive receptors considered particularly vulnerable was also conducted, including (elementary/secondary) schools, daycares, city playgrounds, and residential elderly care facilities. The locations of these facilities within the SoCAB area were mapped, and multiple radial buffers were created around the sites at the same distances as stimulated wells dataset (100 m (328 feet), 400 m (1,312 feet), 800 m (2,625 feet), 1,000 m (3,281 feet), 1,600 m (5,249 feet), and 2,000 m (6,562 feet)). The counts of stimulated wells were then generated for each sensitive receptor at each buffer distance. For pools <50% stimulated, the count of all active production wells falling within the buffer area was generated, and the “well stimulation probability” for that particular pool was then applied.

This assessment focuses on public health risks associated with potential air-pollutant exposures from all active wells and the fraction of these active wells that are hydraulically fractured. Central to that the focus on actively or potentially stimulated wells is the assumption that “plugged,” “buried,” or “idle,” and “abandoned” wells that are not producing according to DOGGR records are not emitting air pollutants. While studies suggest that emissions of methane, VOCs and other TACs from plugged, idle, abandoned and orphaned wells may not be negligible (Dessault et al., 2014; Kang et al., 2014) far more emissions are sourced from wells and ancillary infrastructure while currently active (Thompson et al., 2014), and as such our approach is conservative.

4.B.7. Approach to Analyzed Distances from Stimulated Wells

We conducted a proximity analysis to understand the numbers, densities, and demographics of residents living at various distances from oil and gas wells that have been or are currently being hydraulically fractured and are known to be acidized. We also report results on the proximity of sensitive populations, namely children and the elderly. As discussed in Volume II, Chapter 6, the distance of 800 m (½ mile) and increased density of active oil and gas wells—stimulated and developed with other methods—has been documented in studies in Colorado to increase risks of exposure to multiple TACs that are associated with cancers, respiratory diseases, and birth defects (McKenzie et al., 2012; McKenzie et al. 2014). A U.S. EPA report on dilution of conserved TACs indicates that the dilution at 800 m (½ mile) is on the order of 0.1 mg/m^3 per g/s (U.S. EPA, 1992). Going out to 2,000 m (~1.25 mi) increases this dilution to 0.015 mg/m^3 per g/s. Given that there is increased risk of exposures to benzene at a dilution of 0.1 mg/m^3 per g/s there are concerns that unsafe exposures beyond 800 m may occur. Many conserved TACs known to be associated with oil and gas development (e.g., benzene) persist in the atmosphere long enough to be transported to 2,000 m (~1.25 mi), which is the furthest distance that we assess in our geospatial analysis. As such, we report on human populations within distances from active oil and gas wells, and active wells that have been stimulated, which have been observed in the literature to be potentially hazardous (800 m or $<1/2$ mile), and we also analyze other distances up to 2,000 m (~1.25 mi), given that a number of the known TACs emitted from oil and gas operations in California (and these operations in general) persist in the atmosphere long enough to potentially be atmospherically transported to these distances (e.g., benzene).

4.B.8. References

California Department of Education. School Enrollment, 2014. DataQuest. http://data1.cde.ca.gov/dataquest/. Accessed 10/1/14

California Department of Social Services, 2014. Residential Elderly Care Facilities. https://secure.dss.cahwnet.gov/ccld/securenet/ccld_search/ccld_search.aspx. Accessed 12/1/14

Desseault, M, and R. Jackson, 2014, Seepage Pathway Assessment for Natural Gas to Shallow Groundwater during Well Stimulation, Production and after Abandonment. GEOMontreal. http://shale.palwv.org/wp-content/uploads/2014/02/Dusseault-Jackson-2013-Paper-250_GeoMontreal-Leaking-wells5.pdf. Accessed October 10, 2014.

GreenInfo Network, June 2012. California Recreational Facilities. www.parkinfo.org/recreation. Accessed 10/1/14

Kang, M, C.M. Kanno, M.C. Reid, X. Zhang, D.L. Mauserall, M.A. Celia, Y. Chen, and T.C. Onstott, 2014. Direct Measurements of Methane Emissions from Abandoned Oil and Gas Wells in Pennsylvania. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 111 (51), 18173–18177; doi: 10.1073/pnas. 140831511.

McKenzie LM, Witter RZ, Newman LS, Adgate JL., 2012. Human Health Risk Assessment of Air Emissions from Development of Unconventional Natural Gas Resources. Sci. Total Environ., 424, 79–87; doi:10.1016/j.scitotenv.2012.02.018.

McKenzie, L.M., R. Guo, R.Z. Witter, D.A. Savitz, L.S. Newman, and J.L. Adgate, 2014. Birth Outcomes and Maternal Residential Proximity to Natural Gas Development in Rural Colorado. Environmental Health Perspectives 122; doi:10.1289/ehp.1306722.

State of California, 2014. Geoportal. http://portal.gis.ca.gov/geoportal/catalog/search/resource/details.page?uuid={2BA6C6BD-2803-4CA3-BB10-EF914B96B4A2}. Accessed 12/1/14.

Thompson, C.R., J. Hueber, and D. Helmig, 2014. Influence of Oil and Gas Emissions on Ambient Atmospheric Non-methane Hydrocarbons in Residential Areas of Northeastern Colorado. Elementa: Science of the Anthropocene, 2, 000035; doi:10.12952/journal.elementa.000035.

US Census Bureau, 2010. 2010 Decennial Data. https://www.census.gov/mp/www/cat/decennial_census_2010/summary_file_1_1.html. Accessed 12/1/14.

U.S. Census Bureau, 2012. Census Block SF1 data. http://www2.census.gov/census_2010/04-Summary_File_1/. Accessed 12/1/14.

U.S. Census Bureau, 2014. Tiger Line Shapefiles. Index. ftp://ftp2.census.gov/geo/tiger/TIGER2014/. Accessed 10/1/14.

U.S. Census Bureau, 2014. http://factfinder.census.gov/faces/nav/jsf/pages/download_center.xhtml. Accessed 12/1/14.

U.S. EPA, 1992. Screening Procedures for Estimating the Air Quality Impact of Stationary Sources, Revised. Available at: http://www.epa.gov/oppt/exposure/presentations/efast/usepa_1992b_sp_for_estim_aqi_of_ss.pdf