

Identification and characterization of high methane-emitting abandoned oil and gas wells

Mary Kang^{a,1}, Shanna Christian^b, Michael A. Celia^c, Denise L. Mauzerall^{c,d}, Markus Bill^e, Alana R. Miller^c, Yuheng Chen^b, Mark E. Conrad^e, Thomas H. Darrah^f, and Robert B. Jackson^{a,g}

^aEarth System Science, Stanford University, Stanford, CA 94305; ^bGeosciences, Princeton University, Princeton, NJ 08544; ^cCivil and Environmental Engineering, Princeton University, Princeton, NJ 08544; ^dWoodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08544; ^eEarth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; ^fDivision of Solid Earth Dynamics and Water, Climate, and the Environment, Ohio State University, Columbus, OH 43210; and ^gWoods Institute for the Environment and Precourt Institute for Energy, Stanford University, Stanford, CA 94305

Edited by Steve W. Pacala, Princeton University, Princeton, NJ, and approved September 23, 2016 (received for review April 19, 2016)

Recent measurements of methane emissions from abandoned oil/gas wells show that these wells can be a substantial source of methane to the atmosphere, particularly from a small proportion of high-emitting wells. However, identifying high emitters remains a challenge. We couple 163 well measurements of methane flow rates; ethane, propane, and *n*-butane concentrations; isotopes of methane; and noble gas concentrations from 88 wells in Pennsylvania with synthesized data from historical documents, field investigations, and state databases. Using our databases, we (i) improve estimates of the number of abandoned wells in Pennsylvania; (ii) characterize key attributes that accompany high emitters, including depth, type, plugging status, and coal area designation; and (iii) estimate attribute-specific and overall methane emissions from abandoned wells. High emitters are best predicted as unplugged gas wells and plugged/vented gas wells in coal areas and appear to be unrelated to the presence of underground natural gas storage areas or unconventional oil/gas production. Repeat measurements over 2 years show that flow rates of high emitters are sustained through time. Our attribute-based methane emission data and our comprehensive estimate of 470,000–750,000 abandoned wells in Pennsylvania result in estimated state-wide emissions of 0.04–0.07 Mt (10^{12} g) CH₄ per year. This estimate represents 5–8% of annual anthropogenic methane emissions in Pennsylvania. Our methodology combining new field measurements with data mining of previously unavailable well attributes and numbers of wells can be used to improve methane emission estimates and prioritize cost-effective mitigation strategies for Pennsylvania and beyond.

abandoned wells | oil and gas development | methane emissions | high emitters | climate change

Methane is a potent greenhouse gas (GHG) with a global warming potential 86 times greater than carbon dioxide over a 20-y time horizon (1). A reduction of methane emissions can lead to substantial climate benefits, especially in the short term (2). Recent measurements of methane emissions from abandoned oil and gas wells in Pennsylvania indicate that these wells may be a significant source of methane to the atmosphere (3). Across the United States, the number of abandoned oil/gas wells is estimated to be 3 million or more (4, 5), and this number will continue to increase in the future. As of February of 2016, abandoned oil/gas wells remain outside of GHG emissions inventories, despite evidence that emissions may be substantial nationally. As interest in mitigation of GHG emissions increases, quantifying persistent and large emissions and mitigating them will be increasingly important.

Methane emissions from abandoned wells, as with other fugitive sources in the oil and gas sector, appear to be governed by relatively few high emitters (3, 6–8). It is important for current and future abandoned wells to identify the characteristics that lead to high emissions. This information can provide a rationale for prioritized mitigation.

The century-and-a-half-long history of oil and gas development in Pennsylvania and other US states, such as Texas and California,

has resulted in millions of abandoned wells, and in many cases, poorly documented or missing well records (3, 9–11). As a result, there is a lack of data to characterize abandoned oil and gas wells and the possible relationship between methane emissions and well attributes. Well attributes that may be correlated with methane emissions include depth, plugging status, well type, age, wellbore deviation, geographic location, oil/gas production, and abandonment method (9, 10, 12–14). Previous studies have been limited to wells and attributes with readily available data (12, 14). However, compilation and analysis of historical documents, modern digital databases, and field investigations can be used to infer well attributes of the many wells without data. In this work, we focus on Pennsylvania, which has the longest history of oil and gas development, to determine and explore the role of well attributes, mainly depth, plugging status, well type (e.g., gas or oil), and coal area designation as well as proximity to subsurface-based energy activities, on methane leakage.

Previously estimated numbers of abandoned wells in Pennsylvania range from 300,000 to 500,000 (3, 15) and are based on either incomplete databases or qualitative expert opinion. The Pennsylvania Department of Environmental Protection (DEP) manages oil and gas well data and has records of only 31,676 abandoned oil and gas wells for the state as of October of 2015. Only 5% of the wells in Pennsylvania measured in an earlier study (3) were on the DEP's list. Furthermore, because of changes in governing bodies and regulations over time, the quality of available records is likely to be poorer for older wells (15). To estimate the actual number of wells,

Significance

Millions of abandoned oil and gas wells exist across the United States and around the world. Our study analyzes historical and new field datasets to quantify the number of abandoned wells in Pennsylvania, individual and cumulative methane emissions, and the attributes that help explain these emissions. We show that (i) methane emissions from abandoned wells persist over multiple years and likely decades, (ii) high emitters appear to be unplugged gas wells and plugged/vented gas wells, as required in coal areas, and (iii) the number of abandoned wells may be as high as 750,000 in Pennsylvania alone. Knowing the attributes of high emitters will lead to cost-effective mitigation strategies that target high methane-emitting wells.

Author contributions: M.K. designed research; M.K., S.C., M.A.C., D.L.M., M.B., A.R.M., Y.C., M.E.C., T.H.D., and R.B.J. performed research; M.K., S.C., M.B., and T.H.D. contributed new reagents/analytic tools; M.K., S.C., M.B., A.R.M., Y.C., and T.H.D. analyzed data; and M.K., S.C., M.A.C., D.L.M., M.B., Y.C., M.E.C., T.H.D., and R.B.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

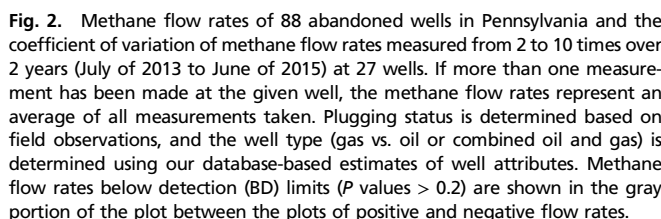
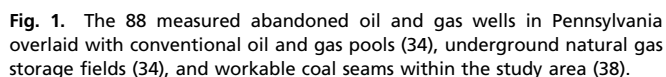
Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. Email: cm1kang@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1605913113/-DCSupplemental.

To identify and characterize high methane-emitting abandoned oil/gas wells, we provide in this paper (i) a database of previously unavailable attributes of measured abandoned wells; (ii) 122 additional field measurements over multiple seasons of methane flow rates and geochemical data, including previously unavailable hydrogen isotopes of methane and noble gas data; (iii) improved estimates of well numbers based on all available data sources; and (iv) an attribute-based methane emissions estimate for abandoned oil and gas wells in Pennsylvania. These data and the associated analysis framework will improve estimates of methane emissions from abandoned oil and gas wells and help develop mitigation strategies across Pennsylvania and beyond.

Methane Flow Rates and Well Attributes. Methane flow rates span from below detection (BD) to 10^5 mg h⁻¹ well⁻¹ for positive methane flow rates (sources of methane to the atmosphere) and from BD to -10^1 mg h⁻¹ well⁻¹ for negative methane flow rates (sinks of methane from the atmosphere) (Fig. 2). Most methane flow rates from abandoned wells (90%) are positive, and all negative numbers are small in magnitude.



Methane flow rates are measured from different categories of abandoned wells in Pennsylvania. For the measured wells without well records, plugging status is determined based on field observations, and the well type (gas vs. oil or combined oil and gas) is determined based on our estimates of well attributes from our assembled database (*SI Appendix*). Across the dataset, abandoned gas wells, specifically unplugged and plugged/vented wells (Pennsylvania Code, Chapter 78), have the highest observed rates of methane emissions (Fig. 2). Abandoned oil wells have consistently lower emissions compared with abandoned gas wells (Fig. 2). The highest measured methane flow rate is $3.5 \times 10^5 \text{ mg h}^{-1} \text{ well}^{-1}$ at an unplugged gas well in McKean County, and the second highest is $2.9 \times 10^5 \text{ mg h}^{-1} \text{ well}^{-1}$ at a plugged but vented gas well in Clearfield County. Venting of plugged wells is required in coal areas, which in Pennsylvania, include regions where mineable coal seams exist (*SI Appendix*).

Methane flow rates are most strongly related to well type (W ; gas vs. oil or combined oil and gas), plugging status (P), and coal area designation (C) (Table 1 and *SI Appendix, Table S3*). No strong trends are observed between methane flow rates and well depth (d), distance to the nearest unconventional well (r_U), or distance to the nearest underground natural gas storage field (r_S). A multilinear fit of d , W , P , C , r_U , and r_S to $\ln \dot{m}$, where $\dot{m}$ ($\text{mg hour}^{-1} \text{ well}^{-1}$) is the methane flow rate, gives an R^2 value of 0.44 and a P value of 4.4×10^{-8} . The P values for the intercept, C , P , and W are below 0.05 and range from 2×10^{-6} (for C) to 0.04 (for the intercept). The P values for d , r_S , and r_U are high at 0.3, 0.8, and 0.4, respectively. The statistically significant well attributes (P values < 0.05) based on the multilinear regression analysis (Table 1 and *SI Appendix, Table S3*) are used in methane emissions estimation. The methane emission factors for nine well categories defined by combinations of W , P , and C range from 1.2×10^{-2} to $6.0 \times 10^4 \text{ mg h}^{-1} \text{ well}^{-1}$ (Table 2).

Methane Flow Rates over Time. Repeat measurements of the same abandoned wells conducted 2–10 times (July of 2013 to June of 2015) (*SI Appendix, Table S2*) show that high emitters ($\geq 10^4$ mg h⁻¹ well⁻¹) have relatively low coefficients of variation, with values ranging from 0.04 to 0.3 (Fig. 2). This result implies that high emitters are emitting methane at consistent levels over multiple years. The coefficient of variation decreases with increasing methane flow rates, implying that lower emitters are more likely to be influenced by variable factors, such as seasonal impacts and measurement error. We also find that the coefficient of variation is

Table 1. Variable coefficients of the multilinear model with R^2 value of 0.44 and P value of 4.4×10^{-8}

Variable in model	Variable coefficient
Intercept	2.84*
D	0.00039
C = coal area	-5.50***
P = unplugged	3.99***
P = plugged/vented	8.33***
W = oil	-2.88*
r_s	0.016
r_u	-0.087

These results are for model L6b in *SI Appendix*, Table S3. The results of additional models are shown and discussed in *SI Appendix*. P values are noted (* $P < 0.05$; *** $P < 0.001$).

unrelated to the number of repeat measurements (*SI Appendix*, Fig. S2).

Geochemistry. The origin of methane from high-emitting wells is predominantly thermogenic, with $\delta^{13}\text{C-CH}_4$ values ranging from -33 to -45‰ (Fig. 3). [Thermogenic methane typically has $\delta^{13}\text{C-CH}_4$ values greater than ~ -40 to -50 ‰, whereas microbial methane typically has $\delta^{13}\text{C-CH}_4$ values below -50 ‰ (17, 19, 20); intermediate $\delta^{13}\text{C-CH}_4$ values, around -50 ‰, can represent mixed thermogenic and microbial sources.] The ratio of $\text{C}_{2-4}/\text{C}_1$ confirms the thermogenic source of high emitters, because the ratio ranges from 0.01 to 0.2. [Microbial sources of methane typically have ratios less than 0.0005 (19, 21).] A larger range in both $\delta^{13}\text{C-CH}_4$ and $\text{C}_{2-4}/\text{C}_1$ values is observed for oil compared with gas wells, with oil wells more likely to emit methane in the microbial range. We do not observe a strong difference in methane isotopes or hydrocarbon ratios between plugged and unplugged wells, although we find that plugged/vented wells have narrower ranges in $\delta^{13}\text{C-CH}_4$ and $\text{C}_{2-4}/\text{C}_1$ values. Wells in coal areas tend to have lower $\text{C}_{2-4}/\text{C}_1$ ratios, regardless of their plugging status or well type, with ratios ranging from 0.001 to 0.04. For wells (in any area) where both $\delta^{13}\text{C-CH}_4$ and $\delta^2\text{H-CH}_4$ are analyzed, most are found to be within the thermogenic range for gases associated with oil reservoirs (17).

High methane-emitting gas wells are found to have the following noble gas ratios: $^3\text{He}/^4\text{He} < 0.10R_A$ (where R/R_A is the

ratio of ^3He to ^4He in a sample compared with the ratio of those isotopes in air, and R_A nomenclature denotes the $^3\text{He}/^4\text{He}$ ratios of samples with respect to air), $^4\text{He}/^{22}\text{Ne} > 100$, and $\text{CH}_4/^{36}\text{Ar} > 1,000$ (Fig. 3 and *SI Appendix*, Fig. S4). ^4He occurs in very low abundances in the atmosphere and is not produced in association with biogenic methane (22). By comparison, ^{22}Ne and ^{36}Ar are ubiquitous, well-mixed, and uniform in the atmosphere. As a result, the noble gases and specifically, elevated levels of ^4He or ratios of thermogenic gases (^4He or CH_4) to atmospheric gases are able to identify high thermogenic methane-emitting gas wells, which cannot always be achieved with hydrocarbon-based geochemical information alone (23).

Number of Abandoned Wells. Using comprehensive databases (15, 24) and analysis of historical documents (25–28) (*SI Appendix*), we estimate the number of abandoned wells in Pennsylvania to be between 470,000 and 750,000 (*SI Appendix*, Table S4). The key difference between our well numbers and previous lower estimates is that we include additional wells drilled for enhanced recovery (ER) purposes (*SI Appendix*). Similar to oil and gas wells used for production, injection wells drilled for water flooding, a widely used enhanced oil recovery technique (26, 29), can also act as pathways for methane and other fluid migration. The data show that the inclusion of ER wells leads to an increase in estimated well numbers by multiplicative factors of 1.7–3.5. We base our estimate of ER wells using these factors for years before 1950, for which the number of ER wells is unknown. There also are discrepancies among the numerous data sources available in historical documents and modern digital datasets (Fig. 4). We compare the data sources to estimate the potential degree of error, which is included as multiplicative factors of 1.3–1.5 in the upper bound estimate (*SI Appendix*, Table S4).

Methane Emission Estimates. The emission factors (Table 2) are combined with the number of wells in each well category in the Pennsylvania DEP database (24) (Fig. 5). The methane emissions contributed by gas wells and wells in coal areas are significantly larger than their share in well numbers. Considering each attribute independently, wells in coal areas represent 21% of the DEP database but 72% of the estimated methane emissions; similarly, gas wells represent 32% of the DEP database but 77% of the methane emissions (Fig. 5). Plugged wells, including those that are vented, represent an estimated 74% of the methane emissions, slightly

Table 2. Emission factors based on coal indicator, plugging status, and well type

Well type and coal area designation	Emission factor ($\text{mg}\cdot\text{h}^{-1}\cdot\text{well}^{-1}$)		No. of measured wells		SE	
	Unplugged	Plugged	Unplugged	Plugged	Unplugged	Plugged
All						
None	2.2×10^4	11.5×10^4	53	35	9.2×10^3	1.0×10^4
Coal	1.2×10^3	4.3×10^4	17	12	9.9×10^2	2.9×10^4
Noncoal	3.1×10^4	4.5×10^2	36	23	1.3×10^4	2.8×10^2
Oil and combined oil and gas						
None	1.9×10^2	3.3×10^2	34	13	9.7×10^1	2.6×10^2
Coal	1.1	1.2×10^{-2}	13	1	9.1×10^{-1}	n/a
Noncoal	3.1×10^2	3.6×10^2	21	12	1.5×10^2	2.8×10^2
Gas						
None	6.0×10^4 *	2.4×10^4	19	22	2.4×10^4	1.6×10^4
Coal	5.2×10^3	$4.7 \times 10^{4*†}$	4	11	3.9×10^3	3.2×10^4
Noncoal	7.5×10^4 *	5.4×10^2	15	11	2.9×10^4	5.1×10^2

The emission factors are averages of mean methane flow rate measurements per well ($\text{mg}\cdot\text{hour}^{-1}\cdot\text{well}^{-1}$). The corresponding numbers of wells and SEs are shown in the next columns. Coal areas are defined here as wells that overlap with one or more workable coal seams. n/a, not applicable.

*The three highest emission factors are shown.

†The measured plugged wells in coal areas are vented as required by regulations.

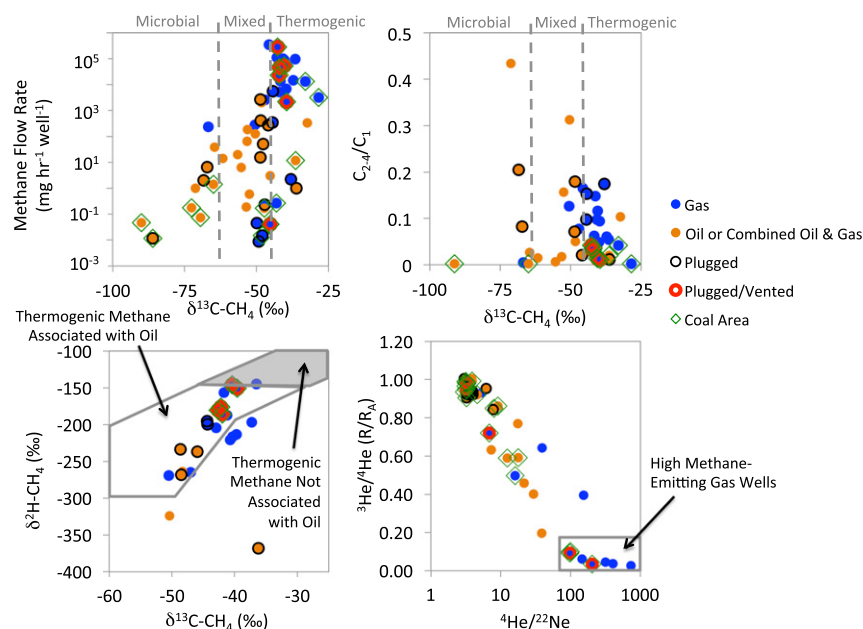


Fig. 3. Carbon and hydrogen isotopes of methane ($\delta^{13}\text{C-CH}_4$ and $\delta^2\text{H-CH}_4$), hydrocarbon concentration ratios ($\text{C}_{2+4}/\text{C}_1$), noble gas data, and methane flow rate data shown colored by well type, circled by plugging status, and marked with green diamond outlines if in a coal area. For repeat measurements, the average of the data for the well is shown. The regions representing thermogenic methane associated and not associated with oil are from ref. 17.

higher than the number for plugged wells (70%) in the DEP database. The DEP database does not distinguish between plugged wells and plugged/vented wells; both are simply categorized as plugged. In our estimate, plugged/vented wells are those that are both plugged and in coal areas, following regulatory requirements in Pennsylvania. Therefore, the methane emissions for all plugged wells (Fig. 5) represent both a large contribution from high methane-emitting plugged/vented gas wells (in coal areas) and a smaller contribution from low methane-emitting plugged wells that are not vented.

Our attribute-based methane emissions estimates for Pennsylvania using improved well numbers range from 0.04 to 0.07 Mt CH_4 per year, which correspond to 5–8% of estimated annual anthropogenic methane emissions for 2011 in Pennsylvania (SI Appendix).

Discussion

Methane Emissions. Well attributes determined for the measured wells in this paper likely remain unavailable for many wells across the United States. Therefore, well attribute estimation studies similar to this analysis may be valuable for many states. For example, West Virginia has at least 57,597 wells that were drilled before 1929 (34% of Pennsylvania wells over the same time period) (25), and records for many wells in the state are likely to be missing. Determining well attributes and numbers is as important as collecting additional measurements for estimating methane emissions. The attributes of high methane-emitting abandoned oil and gas wells identified here as plugging status (P), well type (W), and coal area designation (C) may also be indicative of high emitters elsewhere. In the United States, there are 31 oil-producing states, 33 natural gas-producing states, and 25 coal-producing states (30), with many states simultaneously producing oil, natural gas, and coal. Other well attributes, such as age, wellbore deviation, and operator (12), may also be predictors of methane flow rates. However, we do not explore these attributes here because of a lack of data. Efforts to collect and compile additional well attributes are needed to explore the role of attributes not considered in this study.

The total number of abandoned oil and gas wells remains uncertain in Pennsylvania and across the United States. Documented numbers of wells are more likely to represent lower bounds, because they may not include certain types of wells (e.g., injection wells for ER) and may be missing records. For example, the estimate of 3 million abandoned wells across the United States (4)

does not include injection wells drilled for ER or undocumented wells. In addition, our upper limit in the number of abandoned wells in Pennsylvania of 750,000 may also be an underestimate because of uncertainties associated with differences in terminology among databases and the accuracy of modern digital databases, even in recent records (SI Appendix).

The uncertainties associated with well numbers may be addressed through the application of well-finding technologies (31), field verifications, and database updates. These activities can also help estimate well attributes. In addition, more field measurements of methane emissions are needed from abandoned wells with different attributes and in other geographical locations (i.e., states and countries) to reduce uncertainties in emission factors (32) (Table 2).

Mitigation. Targeting high emitters will lower mitigation costs per unit of methane emissions avoided. The identification of abandoned conventional gas wells and plugged/vented gas wells as the highest emitters allows government agencies to prioritize gas fields and coal areas in their mitigation efforts. Furthermore, explicit categorization of plugged/vented wells, which are found to be high emitters, in state databases may be useful. In addition to database analysis, noble gases, specifically low $^3\text{He}/^4\text{He}$ and high $^4\text{He}/^{22}\text{Ne}$ ratios, provide an independent approach to identify attributes of high methane-emitting abandoned wells.

Because abandoned wells emit methane continuously, over multiple years and presumably many decades, mitigating their emissions will have a larger apparent benefit when longer time periods are considered. Our multiyear measurements show that the high emitters are likely to emit methane at consistently high levels. Such wells may have been emitting at these levels for many decades and will likely continue for decades into the future. A comparison of the benefits of methane emissions reductions from abandoned wells with reductions from intermittent, short-term sources, such as unconventional oil/gas well development, should be performed using emissions integrated over many years.

Well plugging, which is currently viewed as the main mitigation solution (5, 10), does not guarantee a reduction in methane emissions. Plugging was required originally to protect oil and gas reservoirs, reduce risks of explosions, and more recently, protect groundwater. Plugged wells that are vented, as required by regulations in coal areas in Pennsylvania, are very likely to be high

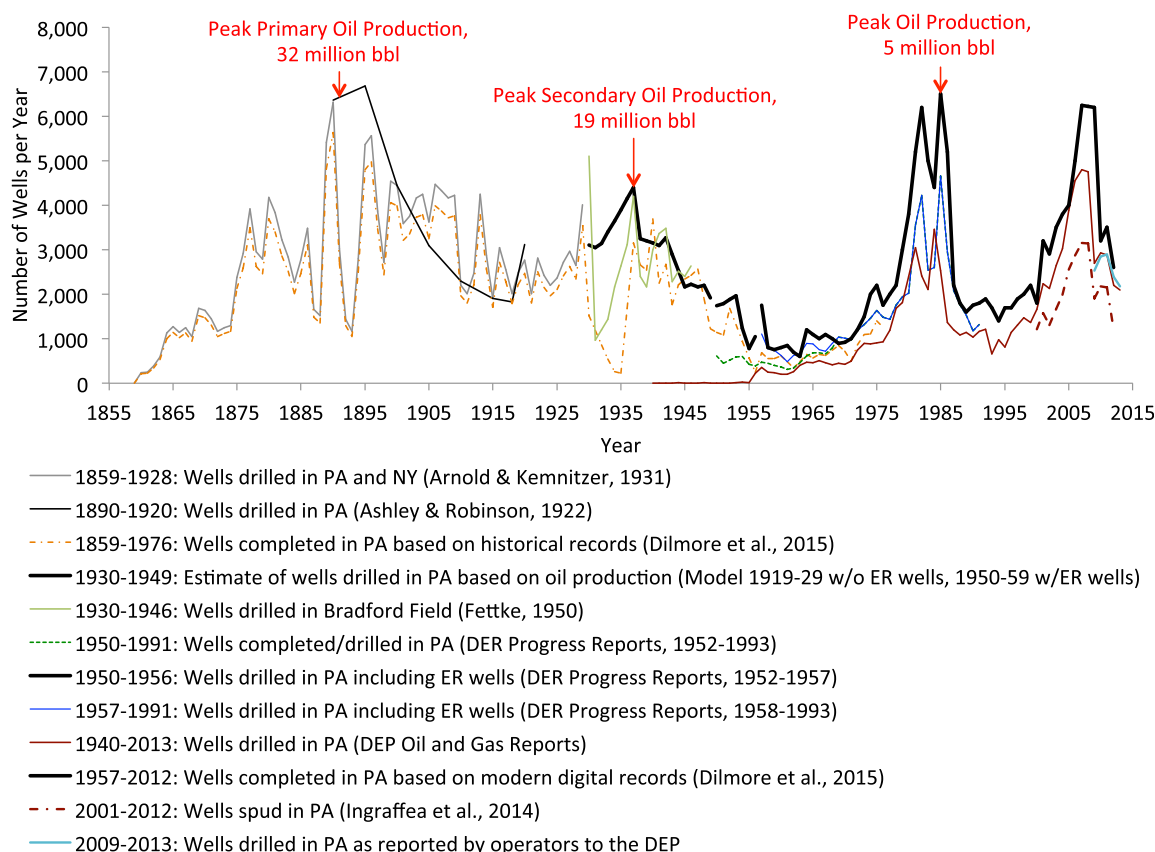


Fig. 4. Number of drilled and/or completed oil and gas wells in Pennsylvania from various historical documents and databases (*SI Appendix*). The thick black lines represent the 1929–2013 data used to estimate the total number of wells (*SI Appendix*, Table S4, second column). For 1859–1928, we use a total well number provided in ref. 25, and the curves shown here are not used to estimate well numbers.

emitters. There are many oil- and gas-producing states with geographically extensive coal layers (e.g., Colorado, Illinois, Indiana, Kentucky, Ohio, Oklahoma, Pennsylvania, West Virginia, and Wyoming). These states have special decommissioning or plugging requirements for coal areas (10). States that require venting in coal areas may want to consider alternatives that ensure safety while reducing methane emissions.

Conclusions

High methane-emitting abandoned wells are found to be unplugged gas wells in noncoal areas and plugged but vented gas wells in coal areas, and they seem to be unrelated to the presence of underground natural gas storage areas or unconventional oil/gas production. The identification of these high emitters provides an opportunity to target mitigation efforts and reduce mitigation costs.

Our attribute-based estimate of 5–8% of estimated annual anthropogenic methane emissions in Pennsylvania is higher than previous estimates, which were based on a single emission factor for all wells and a smaller well count (3, 8, 15). The methane flow rates characterized by well attributes may provide insight into potential emissions outside of Pennsylvania in the 33 oil- and gas-producing US states and other oil- and gas-producing countries. Using the analysis framework presented here, scientists and policymakers can better estimate methane emissions and develop cost-effective mitigation strategies for the millions of abandoned oil and gas wells across the United States and abroad.

Materials and Methods

Well Attributes and Numbers. To determine attributes of the measured wells and estimate the number of abandoned oil and gas wells, we combine information from different types of data sources: historical documents, published literature,

field investigations, and state databases. Historical documents include Pennsylvania agency reports (26–28) and books (25, 33). State databases, including geospatial data, were obtained from the Pennsylvania Department of Conservation and Natural Resources (DCNR) (34) and the Pennsylvania DEP (24), agencies that emerged in 1995 from the Pennsylvania Department of Environmental Resources (DER). We combine and analyze the information to estimate attributes of measured wells based mainly on their location with respect to nearby or overlying oil/gas wells, pools, and fields with attributes in the DCNR database. The attributes determined are depth (d), coal area designation (C), plugging status (P), well type (W), distance to nearest natural gas storage field (r_s), and distance to nearest unconventional oil and gas well (r_u). To estimate the number of abandoned wells, we sum the number of wells drilled annually compiled from multiple sources (15, 24, 25, 27, 28, 33, 35) and subtract the number of active wells from the total (24). We include wells drilled for ER purposes and estimate missing well numbers by scaling available well and production data. We also compare data sources to quantify uncertainties in well numbers. Details on the attribute estimation methodology and the well number estimation are provided in *SI Appendix*, *SI Materials and Methods*.

Field and Laboratory Methods. The measurements of methane flow rates and light hydrocarbon (ethane, propane, and n -butane) concentrations (January, March, and June of 2015 samples) followed methods presented in ref. 3. The measurements were performed across seven counties in Pennsylvania (*SI Appendix*, Table S2). The measurements of methane isotopes were performed at Princeton University (3, 36) and Lawrence Berkeley National Laboratory (LBNL). At LBNL, we also analyzed hydrogen isotopes of methane if concentrations were sufficiently high ($\sim 1,200$ ppmv). For October of 2014 and January, March, and June of 2015, we analyzed the samples for the following noble gases, He, Ne, and Ar, at Ohio State University following methods presented in ref. 22. Additional information on the field sampling and the analysis procedures is provided in *SI Appendix*, *SI Materials and Methods*.

Multilinear Regression. We perform a multilinear regression using the following linear model expressed in Wilkinson notation (37):

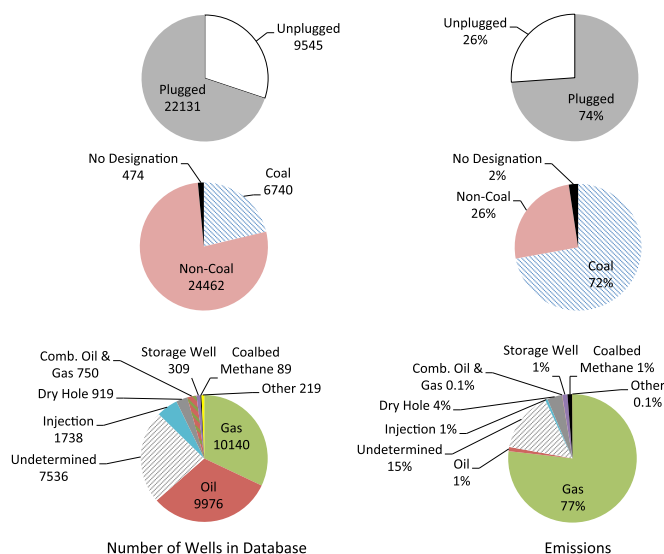


Fig. 5. Number of wells in the PA DEP database (Left) and the corresponding relative methane emissions distribution (Right) based on plugging status, coal area designation, and well type. Each of three attributes is considered independently.

$$\ln \tilde{m} \sim 1 + d + C + P + W + r_5 + r_{10}. \quad [1]$$

Note that the categorical variables, C , P , and W , are denoted using uppercase letters. Multilinear regression is also performed on other linear models, which are summarized in *SI Appendix, SI Materials and Methods*.

Methane Emission Estimates. Based on the multilinear regression results, we use C , P , and W as the key attributes for methane emission estimation:

$$E_{\text{abandoned wells}} = \sum_w \sum_p \sum_c EF_{w,p,c} \cdot n_{w,p,c} \quad [2]$$

where E is the total methane emissions, EF is the emission factor, n is the number of wells, and subscripts w , p , and c represent the appropriate values of W , P , and C , respectively. We consider two well types (w = oil or combined oil & gas and gas), two plugging statuses (P = plugged and unplugged), and two coal area designations (c = coal and noncoal area). We use the Pennsylvania DEP's wells database (24) and the above attributes to determine the proportion of wells in each category. Additional details, including discussions on uncertainties, are given in *SI Appendix, SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank the Stanford Natural Gas Initiative, Princeton University, the Andlinger Center for Energy and the Environment, Stanford University, and the Precourt Institute for Energy. We also thank Venango Senior Environmental Corps (John and Ev Kolojechick, Charlie, and Steve), Clearfield Senior Environmental Corps (Lyle Milland and Rick and Marianne Atkinson), Save Our Streams PA, Joann Parrick, Joe and Cheryl Thomas, Bill Peiffer, Camille Sage Lagron, and Bo Guo for help in the field. For valuable field, laboratory, and planning assistance, we thank David Pal, Ryan Edwards, Ashwin Venkatramen, Matthew Reid, Ejeong Baik, Christianese Kaiser, Eugene Cho, Daniel Ma, Kenneth Campbell, Colin J. Whyte, Myles Moore, and Ben Grove. For assistance in the laboratory and/or helpful comments, we thank Peter Jaffe, Tullis Onstott, Maggie Lau, Tsering W. Shaw, Joseph Vocatur, Harmony Lu, Eric Lebel, Kristin Boye, and Scott Fendorf. Finally, we thank the PA DEP (Stewart Beattie, Scott Perry, Seth Pelepko, and John Quigley) for providing assistance with obtaining data and helpful insights on the data. We acknowledge National Oceanic and Atmospheric Administration Grant NA140AR4310131, the Princeton Environmental Institute, and Vulcan Inc. for supporting this research.

- IPCC (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds Stocker T, et al. (Cambridge Univ Press, Cambridge, UK), p 1535.
- Shindell D, et al. (2012) Simultaneously mitigating near-term climate change and improving human health and food security. *Science* 335(6065):183–189.
- Kang M, et al. (2014) Direct measurements of methane emissions from abandoned oil and gas wells in Pennsylvania. *Proc Natl Acad Sci USA* 111(51):18173–18177.
- Brandt AR, et al. (2014) Energy and environment: Methane leaks from North American natural gas systems. *Science* 343(6172):733–735.
- King GE, Valencia RL (2014) *Environmental Risk and Well Integrity of Plugged and Abandoned Wells, SPE-170949-MS* (Society of Petroleum Engineers, Amsterdam).
- Caulton DR, et al. (2014) Toward a better understanding and quantification of methane emissions from shale gas development. *Proc Natl Acad Sci USA* 111(17):6237–6242.
- Zavala-Araiza D, et al. (2015) Toward a functional definition of methane super-emitters: Application to natural gas production sites. *Environ Sci Technol* 49(13):8167–8174.
- Townsend-Small A, Ferrara TW, Lyon DR, Fries AE, Lamb BK (2016) Emissions of coalbed and natural gas methane from abandoned oil and gas wells in the United States. *Geophys Res Lett* 43(5):2283–2290.
- Jackson RB, et al. (2014) The environmental costs and benefits of fracking. *Annu Rev Environ Resour* 39:327–362.
- Ho J, Krupnick A, McLaughlin K, Munnings C, Shih JS (2016) *Plugging the Gaps in Inactive Well Policy* (Resources for the Future, Washington, DC). Washington, D.C., USA.
- Kang M, Jackson RB (2016) Salinity of deep groundwater in California: Water quantity, quality, and protection. *Proc Natl Acad Sci USA* 113(28):7768–7773.
- Watson TL, Bachu S (2009) *Evaluation of the Potential for Gas and CO₂ Leakage Along Wellbores, SPE 106817* (Society of Petroleum Engineers, Amsterdam).
- Kang M, Baik E, Miller AR, Bandilla KW, Celia MA (2015) Effective permeabilities of abandoned oil and gas wells: Analysis of data from Pennsylvania. *Environ Sci Technol* 49(7):4757–4764.
- Boothroyd IM, Almond S, Qassim SM, Worrall F, Davies RJ (2016) Fugitive emissions of methane from abandoned, decommissioned oil and gas wells. *Sci Total Environ* 547:461–469.
- Dilmore RM, Sams JL, 3rd, Glosser D, Carter KM, Bain DJ (2015) Spatial and temporal characteristics of historical oil and gas wells in Pennsylvania: Implications for new shale gas resources. *Environ Sci Technol* 49(20):12015–12023.
- Darrah TH, Vengosh A, Jackson RB, Warner NR, Poreda RJ (2014) Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales. *Proc Natl Acad Sci USA* 111(39):14076–14081.
- Schoell M (1988) Multiple origins of methane in the earth. *Chem Geol* 71:1–10.
- Jackson RB, et al. (2013) Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proc Natl Acad Sci USA* 110(28):11250–11255.
- Schoell M (1980) The hydrogen and carbon isotopic composition of methane from natural gases of various origins. *Geochim Cosmochim Acta* 44:649–661.
- Jenden P, Drazan D, Kaplan I (1993) Mixing of thermogenic natural gases in northern Appalachian Basin. *Am Assoc Pet Geol Bull* 77:980–998.
- Bernard B, Brooks J, Sackett W (1976) Natural gas seepage in the Gulf of Mexico. *Earth Planet Sci Lett* 31:48–54.
- Darrah TH, et al. (2015) The evolution of Devonian hydrocarbon gases in shallow aquifers of the northern Appalachian Basin: Insights from integrating noble gas and hydrocarbon geochemistry. *Geochim Cosmochim Acta* 170:321–355.
- Hunt AG, Darrah TH, Poreda RJ (2012) Determining the source and genetic fingerprint of natural gases using noble gas geochemistry: A northern Appalachian basin case study. *Am Assoc Pet Geol Bull* 96:1785–1811.
- PA DEP (2015) *Oil and Gas Reports*. Available at www.dep.pa.gov/Pages/default.aspx. Accessed October 16, 2015.
- Arnold R, Kemnitz WJ (1931) *Petroleum in the United States and Possessions* (Harper and Brothers, New York).
- Fettke CR (1950) *Water Flooding in Pennsylvania* (Pennsylvania Department of Internal Affairs, Topographic and Geological Survey), Pennsylvania Geological Survey Fourth Series Bulletin M33; reprinted (1950) (American Petroleum Institute, New York).
- Fettke CR (1951) *Oil and Gas Development in Pennsylvania in 1950* (Pennsylvania Department of Environmental Resources, Pennsylvania Geological Survey), Progress Report 135.
- Cozart CL, Harper JA (1993) *Oil and Gas Development in Pennsylvania* (Pennsylvania Department of Environmental Resources, Pennsylvania Geological Survey), Progress Report 205.
- Lake L (1996) *Enhanced Oil Recovery* (Prentice Hall, Englewood Cliffs, NJ), 1st Ed.
- US Energy Information Administration (2015) *Independent Statistics and Analysis*. Available at www.eia.gov. Accessed December 30, 2015.
- Hammack R, Veloski G, Sams J (2015) *Rapid Methods for Locating Existing Well Penetrations in Unconventional Well Development Areas of Pennsylvania* (San Antonio, TX), URTEC: 2153840.
- CAIT Climate Data Explorer (2015) *Historical Emissions Data* (World Resources Institute, Washington, DC).
- Ashley GH, Robinson JF (1922) *The Oil and Gas Fields of Pennsylvania*, Pennsylvania Geological Survey Fourth Series (Pennsylvania Department of Internal Affairs, Bureau of Topographic and Geological Survey), Vol 1.
- Carter KM, et al. (2015) *Oil and Gas Fields and Pools of Pennsylvania—1859–2011* (PA DCNR, Harrisburg, PA), Open-File Oil and Gas Report 15-01.1.
- Ingraffea AR, Wells MT, Santoro RL, Shonkoff SBC (2014) Assessment and risk analysis of casing and cement impairment in oil and gas wells in Pennsylvania, 2000–2012. *Proc Natl Acad Sci USA* 111(30):10955–10960.
- Chen Y, et al. (2013) Measurement of the 13C/12C of atmospheric CH₄ using near-infrared (NIR) cavity ring-down spectroscopy. *Anal Chem* 85(23):11250–11257.
- Wilkinson GN, Rogers CE (1973) Symbolic description of factorial models for analysis of variance. *J R Stat Soc Ser C Appl Stat* 22:392–399.
- Sholes MA, Skema VW (1974) *Bituminous Coal Resources in Western Pennsylvania* (Pennsylvania Department of Environmental Resources), Mineral Resource Report 68.