Title
A Geospatial Analytical Framework for Understanding Methane Emissions in California

Author
Rafiq, Talha

Publication Date
2022

UNIVERSITY OF CALIFORNIA
RIVERSIDE

A Geospatial Analytical Framework for Understanding Methane Emissions in California

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Environmental Sciences

by

Talha Rafiq

September 2022

Dissertation Committee:
Dr. Francesca Hopkins, Chairperson
Dr. Ying-Hsuan Lin
Dr. William Porter

The Dissertation of Talha Rafiq is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

The Author would like to thank the all the scientists who supported the technical aspects of this work including Dr. Francesca Hopkins, Isis-Frausto Vicencio, Valerie Carranza, Riley Duren, Dr. Charles E. Miller, and Dr. Andrew K. Thorpe.

This work was supported by NASA's Advancing Collaborative Connections for Earth 337 System Science (ACCESS) Methane Source Finder project and by the Lawrence Berkeley National Laboratory under LBNL-CEC agreement PIR-17-015. A portion of the research in this study was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Government support acknowledged. The authors are responsible for the content of the paper and the findings do not represent the views of the funding agencies.

The second chapter of this dissertation, in full, is a reprint of the material as it appears in “Attribution of methane point source emissions using airborne imaging spectroscopy and the Vista-California methane infrastructure dataset” published within *Environmental Research Letters* in 2020 which was authored by Talha Rafiq, Riley M. Duren, Andrew K. Thorpe, Kelsey Foster, Risa Patarsuk, Charles E. Miller, and Francesca Hopkins. The co-author listed in this publication directed and supervised the research which forms the basis for this dissertation.

Dedication

This was a life changing journey for myself as well as my family whose incredible support made this a reality. To my wife, Lauren, thank you for your undying dedication to my success. You were right there when I was excited about starting my Ph.D. You witnessed first-hand all the trials, tribulations, and struggles over these last few years. This work would not have been possible without you.

To my parents, thank you for providing the daily motivation to continue on this path to help the world. The struggles you faced and overcame to place your two sons in these positions in life can never be overstated. I am forever thankful to you both and will continue to dedicate my life to helping our people and our planet. To Hassan, your love and loyalty means everything to me.

I would especially like to thank Dr. Francesca Hopkins for all of her enthusiasm and guidance these past 7 years. From starting out as an intern at NASA JPL to being accepted at UCR in your lab group and having all of our weekly meetings and discussions finally culminate into this research, years later, has been nothing short of worthwhile. Thank you for your patience and belief in helping me develop as a student and as a scientist.

To my friends, Isis, Valerie, and Michelle, thank you for your support and I wish you all success in all your future endeavors. It is my hope you continue to use your talents to help our world.

To the NASA JPL scientists, Riley, Krystal, Chip, and Andrew, thank you all for your guidance during my time there. It was an invaluable experience and one that has undoubtedly shaped the way I approach science.

I finally would like to thank my committee members Dr. Francesca Hopkins, Dr. Ying-Hsuan Lin, and Dr. William Porter for their help and guidance while developing this work as well as UCR's Department of Environmental Sciences for this opportunity.

ABSTRACT OF THE DISSERTATION

A Geospatial Analytical Framework for Understanding Methane Emissions in California

by

Talha Rafiq

Doctor of Philosophy, Graduate Program in Environmental Sciences
University of California, Riverside, September 2022
Dr. Francesca Hopkins, Chairperson

Methane (CH_4), an important greenhouse gas and pollutant, has been targeted for mitigation. Our recent California airborne survey identified >500 CH_4 point sources, which accounted for 34-46% of the statewide CH_4 emissions inventory for 2016. Individual plumes were observed in close proximity to expected CH_4 emitting infrastructure. In order to systematically attribute these plumes to their sources, we developed Vista-CA a geospatial database, that contains more than 900,000 validated CH_4 infrastructure elements in the state of California. In parallel, we developed a complimentary algorithm that attributes any individual CH_4 plume observation to high confidence Vista-CA source with 99% accuracy. This research illustrates the capabilities of the Vista-CA CH_4 database along with Airborne Visible/Infrared Imaging Spectrometer – Next Generation's (AVIRIS-NG) airborne CH_4 retrievals to locate and attribute CH_4 point sources to specific economic sectors. Additionally, this research delivers two emissions products for Kern County: a top-down estimate called the AVIRIS-NG Source Data product and the Vista-CA Bottom-Up emissions dataset. We found general agreement in the source apportionment and magnitudes between these datasets. Moreover, due to current CH_4 inventories having large uncertainties in emissions from the energy processing and production sectors where fugitive emissions predominate, we used airborne CH_4 imaging survey data to show that CH_4 emissions from power plants in California are underestimated by current CH_4 inventory approaches. We developed process-based bottom-up emission estimates for over 300 power plants in California using Intergovernmental Panel on Climate Change (IPCC) methods. We used airborne CH_4 imaging to attribute CH_4 observations to over 250 California power plants and characterize the frequency and persistency of top-down CH_4 emissions. We found that fugitive emissions constitute 90% of total observed emissions from power plants with the remainder derived from process-driven activity while bottom-up emissions are 28 – 54 times smaller than top-down observations. Comparing the inventory-based estimates with observations, the data show "super-emitter" behavior with 60% of total power plant emissions coming from a handful of facilities, likely due to fugitive CH_4 emissions. Future inventories should take advantage of emission observations to quantify CH_4 from these sources to improve the state CH_4 budget and identify mitigation targets.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Methane (CH_4) is the second most important anthropogenic greenhouse gas (GHG) with regards to climate change.[1] Its four hydrogen atoms are held to its single carbon atom by what amounts to a spring. As they float in the air, they bounce around similar to a slinky toy. Pulling one of these atoms far away causes it to snap back together and vice versa. Essentially, CH_4 bounces back and forth with this frequency feature. GHG's like CH_4 and Carbon Dioxide (CO_2) absorb electromagnetic radiation from the Sun which contains all wavelengths of light.[1] The photons from the sun are absorbed by Earth's surface, re-emitted in the infrared, or heat energy part of the spectrum. Thus, a gas which absorbs more of the infrared spectrum of light would be able to trap the most heat. These gases are known as "greenhouse gases". CH_4, due to its physical and chemical makeup, does this work better than CO_2. Thus, being able to absorb more infrared radiation enables it to be a better heat absorber than CO_2, giving it a higher radiative forcing, or heating effect on the Earth system.[1] CH_4 strongly absorbs a few very small wavelength regions occurring around 2.5, 3.5 and 8 microns, all of which are all in the infrared region of the electromagnetic spectrum.[1]

With respect to climate change, CH_4 is a relatively short-lived climate pollutant with respectively around 30 and 85 times the cumulative radiative forcing of CO_2 on a mass basis over 100-year and 20-year lifetimes.[1, 2] Pre-industrialization, global CH_4 atmospheric concentrations leveled around 700 ppb and have since passed 1,900 ppb over the last decade.[1] CH_4 is responsible for ~20% of the radiative forcing since the Industrial Revolution, however, uncertainty in the anthropogenic source distribution of CH_4 emissions presents a challenge for implementing mitigation, globally and in policy relevant domains.[1, 2] A majority of anthropogenic CH_4 sources are separated into two categories: fossil and biogenic. Fossil sources include power plants, refineries, and oil and gas infrastructure while biogenic sources include landfills, dairy livestock farms, rice paddies/wetlands.[1] Atmospheric CH_4 sinks exist in the troposphere in the form of hydroxyl radicals with an average lifetime of around 10 years while in the stratosphere as CH_4 decays when reacting with hydroxyl radicals, chlorine, or $O(^1D)$ over a 120-year lifetime.[1] Surface level sinks include consumption in the soil by methanotrophs, with around 160-year partial lifetime for CH_4.[1] The level of CH_4 continues to rise due to atmospheric and surface level sources primarily caused by anthropogenic activities.[1] The rise in industrial agriculture especially dairy production, increase in the production of waste, and mass adoption of natural gas energy resources has been adding to the global CH_4 budget.

The work presented here focuses on the State of California and investigates its CH_4 emissions budget across relevant sectors. In California, multiple pieces of legislation have been passed to curb this increase in CH_4 while introducing strategies for monitoring and mitigation.[3, 4, 5, 7] In 2006, California passed executive order Assembly Bill 32 titled the Global Warming Solutions Act which required the state to reduce its greenhouse gas emissions levels to 2000 levels by 2010, to 1990 levels by 2020, and to a level 80% below 1990 levels by 2050 and gave power to the California Air Resource Board (CARB) to implement/enforce these standards on industry.[3] In 2015, California passed Assembly Bill 1496 for CH_4 hotpot monitoring which included monitoring and measuring CH_4 hot spots, lifecycle analysis of GHG emissions from natural produced and imported into California, assessing the atmospheric reactivity of CH_4 as a precursor to the formation of photochemical oxidant, and updating policy to incorporate new data.[4] The following year, California passed Senate Bill 32, further defining emissions by mandating reductions in GHG emissions to 40% of 1990 levels by 2030.[5] As a result, total GHG emissions (predominantly CO_2) have been steadily decreasing in California which is a testament to the policies set forth by these legislation, enforcement, and investment bills, but CH_4, on the other hand, has been steadily increasing due to increases in activities and operations from waste, energy, and agriculture sectors.[6] A few examples of these that are covered in this work include California's electric generating sector and Kern County's oil and gas infrastructure. Consequently, in 2016 California enacted legislation in the form of Senate Bill 1383 for specifically targeting the reduction of CH_4.[7] To this end, it is necessary that California has an appropriate and sufficient CH_4 observing system and analytical methodologies to evaluate progress towards its emission reduction goals given large uncertainties in current techniques.

Full accounting of CH_4 from a large geographic area involves utilizing two main methods: top-down observations and bottom-up modeling.[1] Top-down approaches measure patterns through direct atmospheric observation of various CH_4 sources and sinks. Bottom-up approaches require utilizing specific activity data and local processes to develop models that can estimate and scale emissions across sectors. Both top-down and bottom-up methods contain their own advantages and deficiencies. Top-down and bottom-up emission estimates allow for the estimation of CH_4 in a given area and can be used complementary to one another. CH_4 hotspots or point sources are thought to be important to overall emissions and the nature of localized CH_4 emissions enables a geospatial approach towards quantification and assessment.[8, 9, 10] The identification of CH_4 emission sources, both process-based and fugitive, is important for effective development of appropriate policy and corresponding mitigation methods.[1, 6, 8, 9, 10] Process-based emissions result from the operations or processes occurring at a given facility or are the byproduct of combustion activities that are usually persistent given the facility's operational status. Fugitive emissions are the result of incomplete combustion or leaks in the components or infrastructure of a given facility.

However, uncertainty in both the quantification and allocation of California CH_4 emissions hinders mitigation and emission reduction evaluations. Furthermore, a lack of fine-scale CH_4 emission assessments for California necessitates further investigation and improvement of CH_4 emission estimates.[8, 9, 10] Atmospheric top-down measurement estimates of CH_4 emissions in California have shown that bottom-up inventories underestimate total CH_4 emissions by about half in many areas of California.[11, 12, 13, 14, 15] GHG inventories that incorporate information on CH_4 sources often lack necessary spatial information, and inaccurately represent the spatial distribution of these sources. Uncertainties in these inventories limits the performance of inverse modeling and may be the cause of discrepancies in observed top-down and modeled bottom-up estimates with top-down approaches estimating higher emissions than bottom-up methods.[8, 15] Top-down estimates of CH_4 regularly exceed bottom-up estimates due to uncertainties in the apportionment approaches and show that CH_4 emissions for California are underestimated by ~50% or more depending upon the area.[16, 17] Similarly, bottom-up datasets often rely on inadequate sampling outdated emission factors, and/or inaccurate spatial data of facility and site locations.[8, 11, 12, 13, 15] For example, a large fossil signature contribution from energy industries is undercounted in some bottom-up inventories.[10, 18] Additionally, many studies attempting to estimate total emissions lack the ability to resolve their estimates to the sub-facility or component scale as they are too spatially coarse for individual source identification.[17, 19, 20] Moreover, significant differences in CH_4 source apportionment were seen between atmospheric observations and expected emissions for urban areas in California as well. These studies revealed these estimates differed from the source contributions detailed in regionally downscaled versions of the CARB GHG Inventory.

This research provides appropriate context from existing CH_4 regulatory tools from 2 main state and federal repositories: The California Air Resources Board Pollution Mapping Tool (CARB PMT: https://ww3.arb.ca.gov/ei/tools/pollution_map/) and the U.S. Environmental Protection Agency Facility Level Information on GreenHouse gases Tool (EPA FLIGHT: https://ghgdata.epa.gov/ghgp/main.do).[21, 22] CARB and EPA facility emission databases contain data on facilities that exceed annual emissions of 10,000 or 25,000 metric tons CO_2-eq, respectively.[21, 22] It is unclear how accurately CH_4 emissions are represented given that these thresholds for emissions reporting are mostly driven by CO_2 emissions, and CH_4 is more difficult to inventory given the importance of fugitive sources. The primary policy tools are the state and EPA GHG inventories, and facility level information only exists in the CARB PMT and EPA FLIGHT datasets. As such, this research provides a foundation to reconcile these governmental planning inventories (based on Intergovernmental Panel on Climate Change methods) with facility level information.

Top-down measurements have also provided a viable source for CH_4 estimates, especially in California. Precision airborne remote sensing of CH_4 has provided sub-meter spectroscopy of CH_4 point sources over large swaths of geography.[8, 23, 24, 25, 26] This level of detailed information is not available from any publicly available state and federal repositories. This research utilizes this type of airborne data and expands upon work collected as part of the California Methane Survey by NASA Jet Propulsion Laboratory's Next Generation Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-NG). Using AVIRIS-NG, Duren et al. (2019) surveyed CH_4 emission sources in California and visualized CH_4 plumes at the scale of their sources.[8] Their research confirmed the "super-emitter" characteristic where a handful of emission sources constituted a significant amount of the CH_4 budget.[8]

This research provides a foundation for efficient and high-resolution CH_4 measurements over large swaths of geographic area. The individual chapters provide a detailed breakdown of CH_4 emissions analysis using advanced geospatial techniques, geospatial data, and GHG inventory models. In Chapter 2, this works seeks to address how CH_4 point sources be attributed to infrastructure with increased efficiency towards estimation of emissions through geospatial methods.[10] For Chapter 3, we explore how top-down and bottom-up emissions illustrate the impact that oil and gas facilities and infrastructure have on CH_4 emissions in Kern County, California. Finally in Chapter 3, we outline what the impacts of fugitive CH_4 emissions are from California's power plants and their respective super-emitter influence.

I. References

1. National Academies of Sciences, Engineering, and Medicine. 2018. Improving Characterization of Anthropogenic Methane Emissions in the United States. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24987.

2. Etminan, M., G. Myhre, E. J. Highwood, and K. P. Shine. 2016. Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, Geophys. Res. Lett., 43, 12,614–12,623, DOI: doi:10.1002/2016GL071930.

3. Nunez and Pavley. AB-32 Air pollution: greenhouse gases: California Global Warming Solutions Act of 2006. California Legislative Information. 2006. https://leginfo.legislature .ca.gov/faces/billNavClient.xhtml?bill_id=200520060AB32. (accessed 6th September 2022).

4. Thurmond. AB-1496 Methane emissions (2015-2016). California Legislative Information. 2015. https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id=201520160AB1496 (accessed 6th September 2022).

5. Pavley. SB-32 California Global Warming Solutions Act of 2006: emissions limit (2015-2016). California Legislative Information. 2016. https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB32 (accessed 6th September 2022).

6. California Air Resources Board. Methane CH_4 Research Program. 2022. https://ww2.arb.ca.gov/our-work/programs/methane-research (accessed 6th September 2022).

7. Lara. SB-1383 Short-lived climate pollutants: methane emissions: dairy and livestock: organic waste: landfills (2015-2016). 2016. https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB1383 (accessed 6th September 2022).

8. Duren, R. M., Thorpe, A. K., Foster, K.T., Rafiq, T., Hopkins, F. M. Yadav, V., Bue, B. D., Thompson, D. R., Conley, S., Colombi, N.K., Frankenberg, C., McCubbin, I. B., Eastwood, M. L., Falk, M., Herner, J. D., Croes, B. E., Green, R. O., Miller, C. E. 2019. California's Methane 1 Super-Emitters. Nature 575, 180–184. https://doi.org/10.1038/s41586-019-1720-3.

9. Carranza, V., Rafiq, T., Frausto-Vicencio, I., Hopkins, F. M., Verhulst, K. R., Rao, P., Duren, R. M., and Miller, C. E. 2018. Vista-LA: Mapping methane-emitting

infrastructure in the Los Angeles megacity, Earth Syst. Sci. Data, 10, 653-676, DOI: https://doi.org/10.5194/essd-10-653-2018.

10. Rafiq, T. Duren, R. M., Thorpe, A. K., Foster, K., Patarsuk, R., Miller, C. E., Hopkins, F. M. 2020. Attribution of methane point source emissions using airborne imaging spectroscopy and the Vista-California methane infrastructure dataset. Environ. Res., Lett. 15, 1-12, DOI: https://doi.org/10.1088/1748-9326/ab9af8.

11. Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L. 2012. On the Sources of Methane to the Los Angeles Atmosphere. Environmental Science & Technology, 46, 9282-9289, DOI: https://doi.org/10.1021/es301138y.

12. Peischl, J., Ryerson, T. B., Brioude, J., Aikin, K.C., Andrews, A. E., Atlas, E., Blake, D., Daube, B. C., de Gouw, J. A., Dlugokencky, E., Frost, G. J., Gentner, D. R., Gilman, J. B., Goldstein, A. H., Harley, R. A., Holloway, J. S., Kofler, J., Kuster, W. C., Lang, P. M., Novelli, P. C., Santoni, G. W., Trainer, M., Wofsy, S. C., and Parrish, D. D. 2013. Quantifying sources of methane using light alkanes in the Los Angeles basin, California, J. Geophys. Res. Atmos., 118, 4974 – 4990, DOI: https://doi.org/10.1002/jgrd.50413.

13.Cui, Y. Y., Brioude, J., McKeen, S. A., Angevine, W. M., Kim, S.-W., Frost, G. J., Ahmadov, R., Pieschl, J., Bousserez, N., Liu, Z., Ryerson, T. B., Wofsy, S. C., Santoni, G. W., Kort, E. A., Fischer, M. L., and Trainer, M. 2015. Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin, J. Geophys. Res. Atmos., 120, 6698–6711, DOI: https://doi.org/10.1002/2014JD023002.

14. Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., and Blake, D. R. 2014. Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data, J. Geophys. Res. Atmos., 119, 7741–7756, DOI: https://doi.org/10.1002/2014JD021551.

15. Jeong, S., Zhao, C., Andrews, A. E., Bianco, L., Wilczak, J. M., and Fischer, M. L. 2012. Seasonal variation of CH_4 emissions from central California, Journal of Geophysical Research: Atmosphere, 117, DOI: https://doi.org/10.1029/2011JD016896.

16. Alvarez, R. A., Pacala, S. W., Winebrake, J. J., Chameides, W. L., Hamburg. S. P. 2012. Methane leakage from natural gas infrastructure. Proceedings of the National Academy of Sciences, 109, 6435-6440, DOI: http://doi.org/10.1073/pnas.1202407109.

17. Jeong, S., Millstein, D., and Fischer, M. L. 2014. Spatially Explicit Methane Emissions from Petroleum Production and the Natural Gas System in California. Environmental Science & Technology, 48, 5982–5990, DOI: https://doi.org/10.1021/es4046692.

18. Hajny, K. D., Salmon, O. E., Rudek, J., Lyon, D. R., Stuff, A. A., Stirm, B. H., Kaeser, R., Floerchinger, C. R, Conley, S., Smith, M. L., Shepson, P. B. 2019. Observations of Methane Emissions from Natural Gas-Fired Power Plants. Environ. Sci. Technol. 53, 8976 – 8984, DOI: https://doi.org/10.1021/acs.est.9b01875.

19. Maasakkers J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., Hight, C., DeFigueiredo, M., Desai, M., Schmeltz, R., Hockstad, L., Bloom, A. A., Bowman, K. W., Jeong S., and Fischer, M. L. 2016.Gridded National Inventory of U.S. Methane Emissions. Environmental Science & Technology. 50, 13123-13133, DOI: http://doi.org/10.1021/acs.est.6b02878.

20. Yadav, V., Duren, R., Mueller, K.,Verhulst, K. R., Nehrkorn, T., Kim, J., Weiss, R.F., Keeling, R., Sander, S., Fischer, M. L., Newman, S., Falk, Matthias, Kuwayma, T., Hopkins, F., Rafiq, T., Whetstone, J., and Miller, C. 2019. Spatio-temporally resolved methane fluxes from the Los Angeles megacity. Journal of Geophysical Research: Atmosphere, 124, DOI: https://doi.org/10.1029/2018JD030062.

21. California Air Resources Board. 2020. Pollution Mapping Tool Sources in Your Community. https://ww3.arb.ca.gov/ei/tools/pollution_map/pollution_map.htm. (accessed 6th September 2022).

22. Environmental Protection Agency. Facility Level Information on GreenHouse Gases Tool. 2017. https:// ghgdata.epa.gov/ghgp/main.do. (accessed 6th September 2022).

23. Frankenberg, C., Thorpe, A. K., Thompson, D. R., Hulley, G., Kort, E. A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C., Conley, S., Bue, B. D., Aubrey, A. D., Hook, S., Green, R. O. 2016. Airborne methane mapping in the Four Corners area. Proceedings of the National Academy of Sciences. 113 (35) 9734-9739, DOI: 10.1073/pnas.1605617113.

24. Thorpe, A.K., Duren, R., Conley, S., Prasad, K., Bue, B., Yadav, V., Foster, K., Rafiq, T., Hopkins, F., Smith, M., Fischer, M. L., Thompson, D., Frankenberg, C., McCubbin, I., Eastwood, M., Green, R. and Miller, C. E., 2020. Methane emissions from underground gas storage in California. Environ. Res. Letters, DOI: https://doi.org/10.1088/1748-9326/ab751d.

25. Thorpe, A., Frankenberg, C., Aubrey, A.D., Roberts, D., Nottrott, A.A., Rahn, T., Sauer, J.A., Dubey, M., Costigan, K.R., Arata, C., Steffke, A.M., Hills, S., Haselwimmer, C., Charlesworth, D., Funk, C.C., Green, R.O., Lundeen, S.R., Boardman, J.W., Eastwood, M., McFadden, J.P. 2016. Mapping methane concentrations from a controlled release experiment using the next generation airborne visible/infrared imaging spectrometer (AVIRIS-NG). Remote Sensing of Environment. 179. 104-115. 10.1016/j.rse.2016.03.032.

26. Thorpe, A.K., Frankenberg, C., Thompson, D.R., Duren, R.M., Aubrey, A.D., Bue, B.D., Green, R.O., Gerilowski, K., Krings, T., Borchardt, J. and Kort, E.A. 2017. Airborne DOAS retrievals of methane, carbon dioxide, and water vapor concentrations at high spatial resolution: application to AVIRIS-NG. Atmospheric Measurement Techniques, 10(10), 3833-3850, DOI: http://dx.doi.org/10.5194/amt-10-3833-2017.

Chapter 2

Attribution of Methane Point Source Emissions using Airborne Imaging Spectroscopy and the Vista-California Methane Infrastructure Dataset

Rafiq, T. Duren, R. M., Thorpe, A. K., Foster, K., Patarsuk, R., Miller, C. E., Hopkins, F. M. 2020. Attribution of methane point source emissions using airborne imaging spectroscopy and the Vista-California methane infrastructure dataset. Environ. Res., Lett. 15, 1-12, DOI: https://doi.org/10.1088/1748-9326/ab9af8.

I. Introduction

Methane (CH_4) is a powerful greenhouse gas (GHG) responsible for ~20% of radiative forcing since the Industrial Revolution[16]; however, uncertainty in the source apportionment of CH_4 emissions poses a challenge for implementing mitigation, globally and in policy relevant domains. In the State of California, reductions in CH_4 emissions are explicitly required by law (2016 SB 1383) as a way to achieve California's climate goals for significantly reducing overall GHG emissions by 2030.[30, 37] To this end, it is necessary that California has an appropriate and sufficient CH_4 observing system to evaluate progress towards its emission reduction goals given large uncertainties in current techniques.[23]

Notable differences in CH_4 source apportionment have been observed between atmospheric observations and expected GHG emissions for urban areas in the state [20, 38, 40]. These studies used measurements of CH_4 and its tracer species in well-mixed air to infer the contributions of different source sectors to regional CH_4 emissions, and found that these estimates differed from the source contributions detailed in regionally downscaled versions of the California Air Resources Board GHG Inventory (CARB GHG Inventory). Given that the CARB GHG Inventory is the primary tool used for tracking GHG emissions in the state, this discrepancy poses a challenge for verifying state-mandated CH_4 mitigation efforts.

Another policy tool for tracking GHG emissions is facility-level governmental reporting programs. The California Air Resources Board Pollution Mapping Tool (CARB PMT: https://ww3.arb.ca.gov/ei/tools/pollution_map/) and the U.S. Environmental Protection Agency Facility Level Information on GreenHouse gases Tool (EPA FLIGHT:

https://ghgdata.epa.gov/ghgp/main.do) show maps of facilities that exceed annual emissions of 10,000 or 25,000 metric tons CO2-e for CARB and EPA, respectively, along with their annual reported GHG emissions, including CH_4. Facility level emissions tracking is useful because this is the scale where mitigation actions are most often taken; however, these emissions are not verified by independent methods. It is unclear how accurately CH_4 emissions are represented given that these thresholds for emissions reporting are mostly driven by CO_2 emissions, and CH_4 is more difficult to inventory given the importance of fugitive sources and fat-tailed distributions.

Recent advances in airborne remote sensing of CH_4 have enabled meter-scale imaging of CH_4 point sources over areas from 1,000 to 100,000 km^2.[11, 17] Using this technique, Duren et al. (2019) surveyed methane emission sources in California and found that a few hundred CH_4 point sources contributed 34-46% of the overall statewide emissions. The high spatial resolution, 1-3 m per pixel, of airborne imaging spectrometers is capable of visualizing CH_4 plumes at the scale of their sources. Combining plume imagery with detailed geospatial information from high-resolution satellite imagery in a platform such as Google Earth, enables one to attribute CH_4 emissions to specific facilities and infrastructure components.[5, 22, 34] However, analysis of these CH_4 point sources requires an accurate, systematic method to identify infrastructure components at policy relevant (e.g., facility) levels. Such detailed information is not provided in a state level, aggregated inventory such as the CARB GHG Inventory or even in high resolution (~10 km) disaggregated inventories [e.g., Maasakkers et al. 2018].

A different approach for systematically understanding the distribution of CH_4 sources was demonstrated by Carranza et al. (2018) for the Los Angeles (LA) Basin through the development of Vista-Los Angeles (Vista-LA).[4] Vista-LA is a geospatial dataset of all anthropogenic CH_4 infrastructure within the LA Basin that attempts to represent all potential sources of CH_4 emissions regardless of the expected size of emissions. Vista-LA is organized in the same way as the CARB GHG Inventory for sectoral analyses, but is spatially disaggregated with representations of CH_4 emission sources at the facility scale and down to individual components, such as gas pipelines. Combining this detailed dataset with new, high resolution observational data of CH_4 emissions from airborne remote sensing enables a more thorough "inventory" of CH_4 based on actual observations that is likely to be more robust than activity/emission factor methods that do not capture fugitive or anomalously large sources that are thought to be common for CH_4.

In this study, we expand the Vista approach to the whole state of California (Vista-CA) for analysis of CH_4 plume data collected by the Airborne Visible/InfraRed Imaging Spectrometer-Next Generation (AVIRIS-NG) in California in 2016-218. Previously, we showed that Vista-CA was used for survey planning and manual source attribution for a subset of these flights (Duren et al., 2019). Here, we (1) detail further updates to the Vista methodology to enable automated source attribution, and (2) compare its performance to attributions made from existing facility data from government reporting programs,[8, 9] and (3) demonstrate that source attribution can be automated for fast-turnaround data processing for all 2016-2018 plume datasets using the Geospatial Source Attribution Automated Model (GSAAM).

II. Methods

Vista-CA is a geospatial database of 901,009 validated elements of potential CH_4 emitting infrastructure developed from publicly available datasets that have been validated and standardized for the entire state of California (Figure 2.1). Vista-CA includes 17 different CH_4 source layers that have been systematically categorized into facilities and sub-facilities. These include power plants, refineries, natural gas fueling stations, natural gas stations, pipelines, distribution pipelines, natural gas processing plants, natural gas storage fields, oil and gas facilities, oil and gas field boundaries, oil and gas wells, dairies, feedlots, digesters, composting sites, solid waste disposal sites, and wastewater treatment plants (Table 2.1). New source layers were added to Vista-CA that were not present in Vista-LA because they were (1) not included as CH_4 sources in the CARB inventory, but were observed to emit by AVIRIS-NG[8] (composting sites), (2) not present in the LA domain (digesters, feedlots), and (3) comprise datasets that were published after Vista-LA (oil and gas facilities). Federal and state data repositories were used as the primary data sources.[1, 2, 12, 13, 14] These datasets were validated by cross-comparing multiple datasets for spatial consistency and accuracy. We used Google Earth aerial imagery to either identify or confirm geolocations as well to denote geographic extents of individual facilities and infrastructure. All feature datasets were georeferenced and updated with standardized metadata, and are freely available on the Oak Ridge National Laboratory Distributed Active Archive Center for Biogeochemical Dynamics.[21] Vista-LA layers that previously covered only the LA Basin were expanded to include the full extent of California.[4, 9] All datasets are formatted as vectors stored as either lines, points, or polygons (Table S2.1).

Vista-CA is organized according to the CARB GHG Inventory, which itself is based on the framework established by the Intergovernmental Panel on Climate Change (IPCC).[10] However, both IPCC and CARB are process-based inventories that use state level activity data to estimate GHG emissions, whereas Vista-CA is database of actual facilities that may emit CH_4 in California. Organizing Vista-CA source types in this way is critical for comparison with inventory and for categorizing contributions of different emission sectors.

CARB provides annual CH_4 estimates for top-emitting facilities across California.[1] Their pollution mapping tool (PMT) is a geospatial database that enables users to query, locate, and view reported GHG and criteria pollutant emissions at the facility scale.[1] CH_4 data is only reported for facilities emitting >10,000 metric tons CO2e annually. PMT contains facility addresses that are sometimes inaccurate, often giving the address of operator headquarters instead of the emitting facility. True addresses were obtained from publicly available records and were used to validate locations in the CARB PMT data. CARB PMT CH_4 data for 2016 contained a geospatial dataset of 597 CH_4 reporting facilities in California.[1]

EPA FLIGHT is a geospatial database of the locations of approximately 8,000 facilities that report annually to the EPA Greenhouse Gas Reporting Program (GHGRP). EPA FLIGHT tracks facilities that emit more than 25,000 MTCO2eq/year, and accounts for 85-90% of emissions included in the official EPA GHG Reporting Program.[15] EPA

FLIGHT CH_4 data for 2017 contained geospatial data of 389 CH_4 reporting facilities for California.[13] Both CARB PMT and EPA FLIGHT differ from the official GHG inventories from their respective agencies in that they are based on reported emissions at the facility scale, not activity data.

We used CH_4 plume observations from a survey of California conducted using the Airborne Visible/Infrared Imaging Spectrometer - Next Generation (AVIRIS-NG) instrument.[8, 9] AVIRIS-NG is capable of detecting concentrated CH_4 plumes by measuring ground-reflected solar radiation across 427 contiguous spectral bands ranging from 350 to 2,500 nm wavelengths with 5 nm spectral sampling at 3 m spatial resolution. The CH_4 retrieval is based on absorption spectroscopy between 2,100 and 2,500 nm and provides a mixing ratio length that represents CH_4 enhancement integrated along the column beneath the aircraft in parts per million x meter (ppm m).[32] AVIRIS-NG's spectral resolution, high spatial resolution, and high signal-to-noise ratio has permitted high-resolution mapping of CH_4 as well as CO_2 and H_2O.[5, 32, 35, 36] AVIRIS-NG has consistently detected and quantified CH_4 point sources from multiple emissions sectors for emissions as small as 2-10 kg CH_4/hr, depending on surface albedo and aircraft/ground speed.[8, 9] For further information, the specific plume localization and identification process has been detailed by Duren et al. 2019.[8]

We used AVIRIS-NG CH_4 plume observations from 2016-2018 collected during the California Methane Campaign,[8] in which 2,424 CH_4 plumes were identified manually with high confidence. These plume observations are available to the public at the Methane Source Finder web portal (https://data.carbonmapper.org). Emissions from 1,181 of the

2016-2017 plume detections were quantified and published by Duren et al. (2019) along with manual source attribution using Vista-CA. Here we performed source attribution on the additional 748 unpublished plumes from 2016-2017 for which emissions quantification was uncertain, and the 495 plumes from 2018.

The meter-scale resolution and geolocation accuracy of AVIRIS-NG observations enabled us to determine the source location of nearly all CH_4 plumes within a radius of 5 meters or less. We then attributed each plume observation to an emission source facility/sub-facility in the Vista-CA database based on spatial proximity (Figure 2.2). First, we manually identified the emissions origin of each observed CH_4 plume using preliminary versions of Vista-CA. This process entailed overlaying orthorectified grayscale images of CH_4 retrieved by a linearized matched filter (AVIRIS-NG Level 3 data) on high resolution Google Earth aerial imagery for broader context with Vista-CA infrastructure maps simultaneously displayed. This process was conducted for the 1,181 plumes published by Duren et al. (2019). We treat this manual attribution as the true attribution for development of automated attribution algorithms. Next, we automated the attribution of observed AVIRIS-NG CH_4 plumes based on proximity to Vista-CA features (Figure S2.2). We developed a decision-tree framework to attribute AVIRIS-NG plumes to the nearest logical Vista-CA feature while considering the effect of known spatial biases that impact proximity attribution (Figure 2.3). Specifically, there is a large degree of spatial overlap amongst Vista-CA source layers (Table S2.2). These overlaps often occur because Vista-CA layers are organized by source type, e.g., power plants, without considering whether each individual feature is part of a larger facility, such as a landfill or refinery which often contain their own power plants. An overlap analysis was conducted among all 17 Vista-CA source layers (Table S2.2) to distinguish facility vs. sub-facility scale features (Table 2.1). Finally, within each branch and sub-branch of the framework, a specific radius was determined to maximize attribution accuracy and reduce the number of false positives and false negatives (Table 2.1). Using the greatest distance between a methane plume and its source feature in Vista-CA, we determined a radius for each Vista-CA layer using the near function in ArcGIS.

To develop the automated model, we first employed an automated simple distance method to attribute 1,181 plumes from the published Duren et al. dataset to the nearest Vista-CA feature without providing any spatial logic or hierarchal considerations (Figure 2.3). This baseline allowed us to see where improvements would have to be initiated, how to prioritize or develop the data hierarchy, and how to logically assess spatial complexities within the data. Consequently, a hierarchical structure was developed within the decision-tree framework to account for spatial biases in order to reduce the number of misattributions (Figure 2.3). Attributions are done at the facility level, which also gives a sectoral attribution by IPCC source category. Further, we attribute plumes to sub-facilities if present to enable better understanding of the emitting process. For example, if the workflow attributed a given plume to a refinery, it would further assess whether it could also be attributed to relevant sub-facility components such as an oil & gas well or a sub-facility power plant. If so, they would be appropriately attributed; if not, they would simply stay as being attributed to the refinery facility-level. If a plume was unable to be identified by any of these features, then the plume would pass to the next sector for attribution according to the decision tree structure. This process would continue until all sectoral, facility, and sub-facility Vista-CA data is parsed. The remaining un-attributed plumes are labeled “Unknown”.

Manually attributing plumes required significant time and effort; however, the decision-tree workflow was strategically designed for easy automation. The resulting Geospatial Source Attribution Automated Model (GSAAM) is an efficient plume-to-source attribution framework designed with 2 main inputs: latitude/longitude (X, Y) coordinates as a comma-separated values spreadsheet and Vista-CA geospatial datasets. After all attributions have been completed according to the decision tree, the model merges the result together into a final product outputting a tabular spreadsheet along with an ESRI point shapefile (Table S2.1).

III. Results

We used Vista-CA to perform source attribution of 2,424 methane plumes observed by AVIRIS-NG during the 2016-2018 California Methane Survey (Figure 2.2).[8] First, we manually attributed plumes to facilities in Vista-CA to determine the maximum possible number of sources attributed using the Vista-CA dataset. Of the total 2,424 methane plumes observed, 2,407 (99.3%) were manually attributed to a Vista-CA feature (Table 2.2). Unattributed plumes, hereafter called unknowns, were either found far from any methane emitting infrastructure, such as in an agricultural field, or were found associated with methane emission sources not included in Vista-CA, such as a beef processing plant.[8]

We compared manual attribution with the Vista-CA dataset to manual attribution with other CH_4 facility databases: CARB PMT and EPA FLIGHT for a subset of 1,181 plumes with high confidence emissions estimates from 2016-2017 (published in Duren et al., 2019). Manual attribution of airborne CH_4 plume detections with both CARB PMT and EPA FLIGHT data resulted in significantly lower attribution accuracies across the 6 IPCC sectors (Figure 2.4). Use of CARB PMT attributed 39.5% of observed CH_4 plumes (466/1,181), and after comparison with the original manual attribution, only 30.6% of plume attributions using PMT were considered correct (361/1,181 plumes) (Figure 2.4). CH_4 plume attribution with EPA FLIGHT had similar results, with attribution of 38.8% of CH_4 plumes (458/1,181), with 30.9% correct (366/1,181 plumes) (Figure 2.4). Performance of PMT and FLIGHT varied greatly across sectors, with much better source attribution for Energy (IPCC 1A1) and Waste (IPCC 4A1) compared to Oil and Natural Gas (IPCC 1B2) and Manure Management (IPCC 3A2) (Figure 2.4). In total, CARB PMT only had 18% (52) and EPA FLIGHT only had 15% (44) of the 290 unique facilities in Vista-CA that were observed to be emitting CH_4 by AVIRIS-NG in our dataset (Table 2.3).

Next, we used Vista-CA to automate source attribution based on spatial relationships between observed CH_4 plumes and Vista-CA infrastructure. As discussed previously, a simple distance analysis to attribute each CH_4 plume to a Vista-CA feature served as the validation baseline for measuring model performance. Vista-CA GSAAM V4 improved attribution accuracy over the simple distance method from 51.3% to 99.6% at the facility level across all seven IPCC Level 3 source categories. The total number of facilities broken down by Vista categories along with the number of unique correctly attributed facilities across all three datasets was also calculated for a direct comparison of completeness (Table 2.2).

For all plume observations, Vista-CA GSAAM V4 correctly attributed 2,384 of 2,403 (99.2%) total plume observations at the facility level, excluding the 21 plumes from unknown sources described above (Table 2.2). Only 8 plumes were attributed to incorrect Vista-CA facilities, yielding a false positive rate (mis-attributions) of 0.45% (6 times for Harris Ranch Meat Plant and 2 times for the Palos Verdes Landfill). The overall false negative rate, indicating missed attributions when there was in fact a source from manual attribution, was 0.18%. Moreover, GSAAM attributed 19.5% of the plumes to sub-facility level infrastructure with an attribution accuracy of 100% (Table 2.2). We achieved ideal 1:1 plume-to-source attribution accuracies (100%) for three of the six IPCC source categories: 1A1 Energy Industries, 4B Biological Treatment of Solid Waste, and 4D1 & 4D2 Domestic Wastewater Treatment & Discharge, with the other three categories averaging 99.19% (Figure 2.4).

IV. Discussion

The identification, geolocation and attribution of anthropogenic CH_4 emissions remains a major challenge for emissions monitoring and mitigation. We developed a method for high confidence attribution of meter scale CH_4 plume observations to their emission sources at the facility scale by spatially relating the locations of airborne CH_4 plume detections to geographic datasets that represent locations of potential CH_4 emission sources. We demonstrate this using Vista-CA and CH_4 plumes observed by AVIRIS-NG in both a manual and automated mode. We found that the vast majority of CH_4 plumes in California were found in association with infrastructure known to handle or produce CH_4, consistent with the expectation that these large point sources are anthropogenic, and thus potential targets for CH_4 mitigation. Vista-CA and AVIRIS-NG results are visually depicted in NASA JPL's Methane Source Finder (https://data.carbonmapper.org).

We compared the ability of our Vista-CA dataset to attribute CH_4 plume observations to facility-level regulatory datasets, CARB PMT and EPA FLIGHT (Figure 2.4). Unlike GHG emission inventories that encompass emissions at the level of a state, the CARB PMT and EPA FLIGHT reporting program datasets provide facility-level spatial information and reported estimates of CH_4 emissions. However, they were not as effective as Vista-CA for CH_4 source attribution. The threshold for inclusion in CARB PMT and EPA FLIGHT, based on total expected facility GHG emissions, is ill suited for CH_4 emissions that are characterized by fugitive sources and skewed emissions distributions that make inventories of CH_4 challenging to construct. In contrast, Vista-CA was designed to assume that CH_4 emissions can potentially come from any CH_4 relevant infrastructure. Vista-CA includes (1) sources previously omitted from the regulatory inventories, such as composting sites and natural gas fueling stations, (2) sources with emissions expected to be too small or zero, but that might still be emitting, such as closed landfills,[20] and (3) sources for which there are not readily available public maps, such as dairy farms.[19] In addition, Vista-CA has confirmed geolocations for all sources, avoiding the problem of the address of an emitter differing from the actual location of emissions, as occurs in regulatory datasets. Finally, much effort was put into delineating the geographic extents of Vista-CA sources that have large spatial extents, such as landfills. These spatial extents improve the ability of an automated model to match a plume location with its source facility compared to point locations

Proximity-based attribution methods are limited by the availability of datasets that are used to inform them. 1.6% (39/2,424 plumes) of plumes that were either unknown or misattributed, come from 16 sources that are not currently included in Vista-CA: one meat processing plant; one liquified natural gas terminal; two oil and gas tanks that were not associated with an oil and gas field, facility, or refinery; five landfills; two agricultural sites; one dairy; and four related to oil and gas fields with no other spatial details. False negatives—plumes that were not attributed to any feature— persisted mainly due to inconsistent spatial coverage in the oil and gas field boundary dataset. Better accounting for the spatial extents of various facilities in Vista-CA, such as dairies, could reduce these problems, but manual digitizing of facility extents would require significant additional effort. For these more complex or confounding cases, we suggest a "human-in-the-loop" method to reconcile some of these discrepancies.

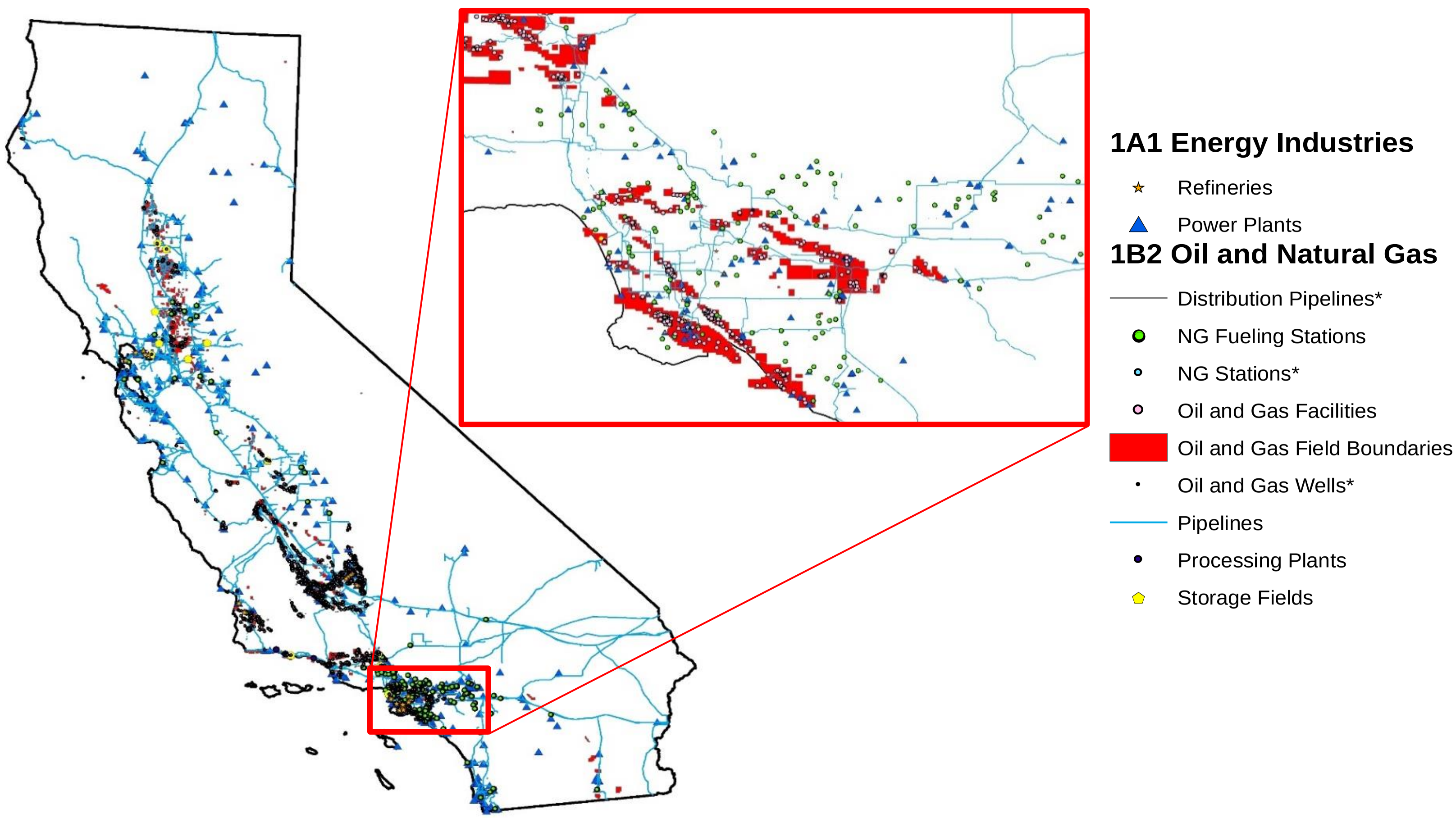
We also distinguished facility level sources from sub-facility features in the Vista-CA dataset to improve automated source attribution. Vista-CA was originally designed with a focus on the facility level because of its relevance for mitigation activities; however, linking CH_4 plume observations to sub-facility level infrastructure can give deeper insight into the process producing emissions. This has been demonstrated with AVIRIS-NG data for underground storage fields and landfills.[5, 34] We recognize that sub-facility level infrastructure included in Vista-CA is very limited, given that we rely on public databases for our data sources. This problem is most acute in oil and gas fields, which account for 122 plume attributions without more detailed sub-facility attribution. Oil fields such as Midway-Sunset can span hundreds of kilometers, but we have limited information on the oil and gas production infrastructure located therein, such as gathering pipelines, storage tanks, and other oil and gas facilities that are present but not currently included in the Vista-CA oil and gas facilities source layer. This can be further improved with more complete accounting of oil and gas production structures located within these CH_4 source areas. Because of the vast extent of oil and gas fields, we include oil and gas fields at the end of the attribution tree to avoid mis-attribution of CH_4 plumes located there to the oil and gas field when another possible CH_4 emission infrastructure is present (e.g., a dairy located on an oil and gas field). In addition, we distinguish urban from non-urban oil and gas fields, since urban oil and gas fields are much more likely to include CH_4 emission sources that are not related to oil and gas production activities.

One assumption of our approach is that the apparent origin of the plume in hyperspectral imagery is indeed its source; this may not be the case under swirling or still wind conditions.[28] This uncertainty is particularly relevant in areas densely populated with potential sources where Vista-CA facilities overlap one another or are in close spatial proximity. In industrial urban areas, for example, high spatial density of sources from multiple collocated sectors complicates source attribution.

By cataloguing all potential CH_4 emission sources in Vista-CA, we add to a growing body of evidence that a small number of emitters contribute to a large fraction of the total CH_4 emissions[8], with 3.3% (290/8,878) of Vista-CA facilities responsible for all CH_4 plumes observed by AVIRIS-NG (Figure 2.2, Table 2.3). Given that the spatial extent of Vista-CA is only 3.46% of California's area, both the Vista-CA spatial model and the attributions of AVIRIS-NG observations to a subset of Vista-CA allows for a more focused approach when it comes to developing mitigation strategies.

Our source attribution methodology can attribute observed CH_4 plumes down to individual sub-facility infrastructure elements, enabling detailed investigation of sectoral contributions of CH_4 point source emitters, comparison to reported emissions at the facility level, reporting of anomalous activity to facility operators, and investigation of emissions distributions within a source category, as demonstrated in Duren et al. (2019). Moreover, we demonstrated that source attribution can be automated, enabling rapid analysis of large surveys. This is a critical step toward operationalizing airborne CH_4 emissions monitoring, and similar approaches may be needed for analyzing CH_4 point sources detected globally by new satellite missions.[7] A typical plume dataset from an airborne campaign consists of 2,000 plumes, and requires roughly 15-20 hours for manual attribution analysis with a tool like Vista in hand, which is reduced to approximately 5 minutes with automation, and close to 99% attribution accuracy (Figure 2.4). While presently limited to the state of California, Vista-CA and GSAAM are useful tools for any future CH_4 monitoring the state undertakes by allowing a more focused mitigation approach. We suggest the Vista approach may also be applied more broadly for CH_4 point source attribution with new imaging spectrometry from airborne and spaceborne platforms. Expanding Vista globally will require additional automation and methods to deal with the different degrees of sectoral data and metadata available in different regions.

Figure 2.1 A.) IPCC Level 1. Energy Vista-CA Maps. Inset map shows the Energy facilities and infrastructure located around the Los Angeles metropolitan area. This map is grouped by the IPCC Level 1 structure. Each IPCC Level 1 map displays all the relevant infrastructure organized within IPCC Level 3 (Table 2.1).



*Natural Gas Stations, Distribution Pipelines, and Oil and Gas Wells are not shown.

Figure 2.1 B) IPCC Level 3. Agriculture, Forestry & Other Land Use Vista-CA Maps. This map is grouped by the IPCC Level 1 structure. Each IPCC Level 1 map displays all the relevant infrastructure organized within IPCC Level 3 (Table 2.1).

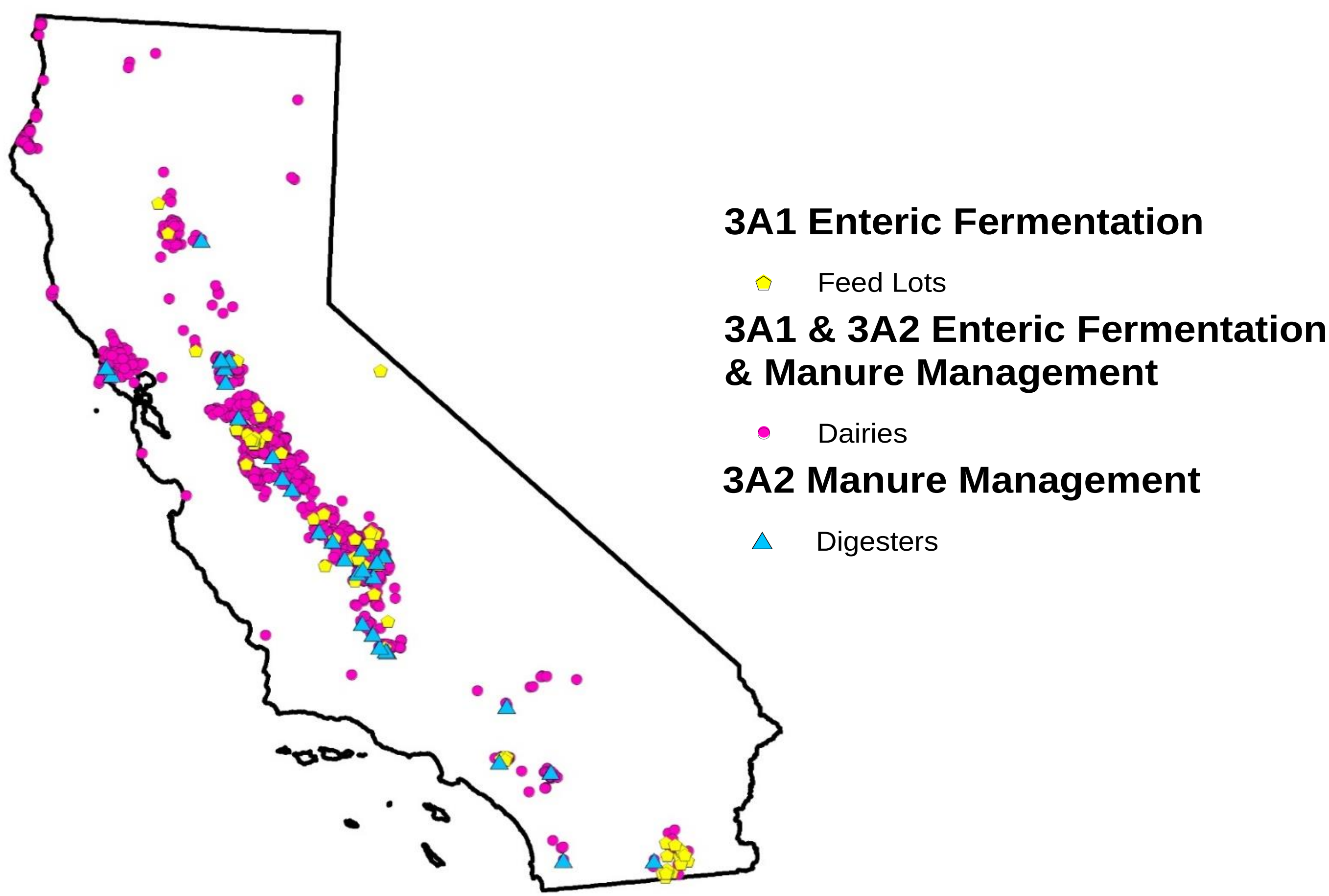


Figure 2.1 C) IPCC Level 4. Waste Vista-CA Maps. This map is grouped by the IPCC Level 1 structure. Each IPCC Level 1 map displays all the relevant infrastructure organized within IPCC Level 3 (Table 2.1).

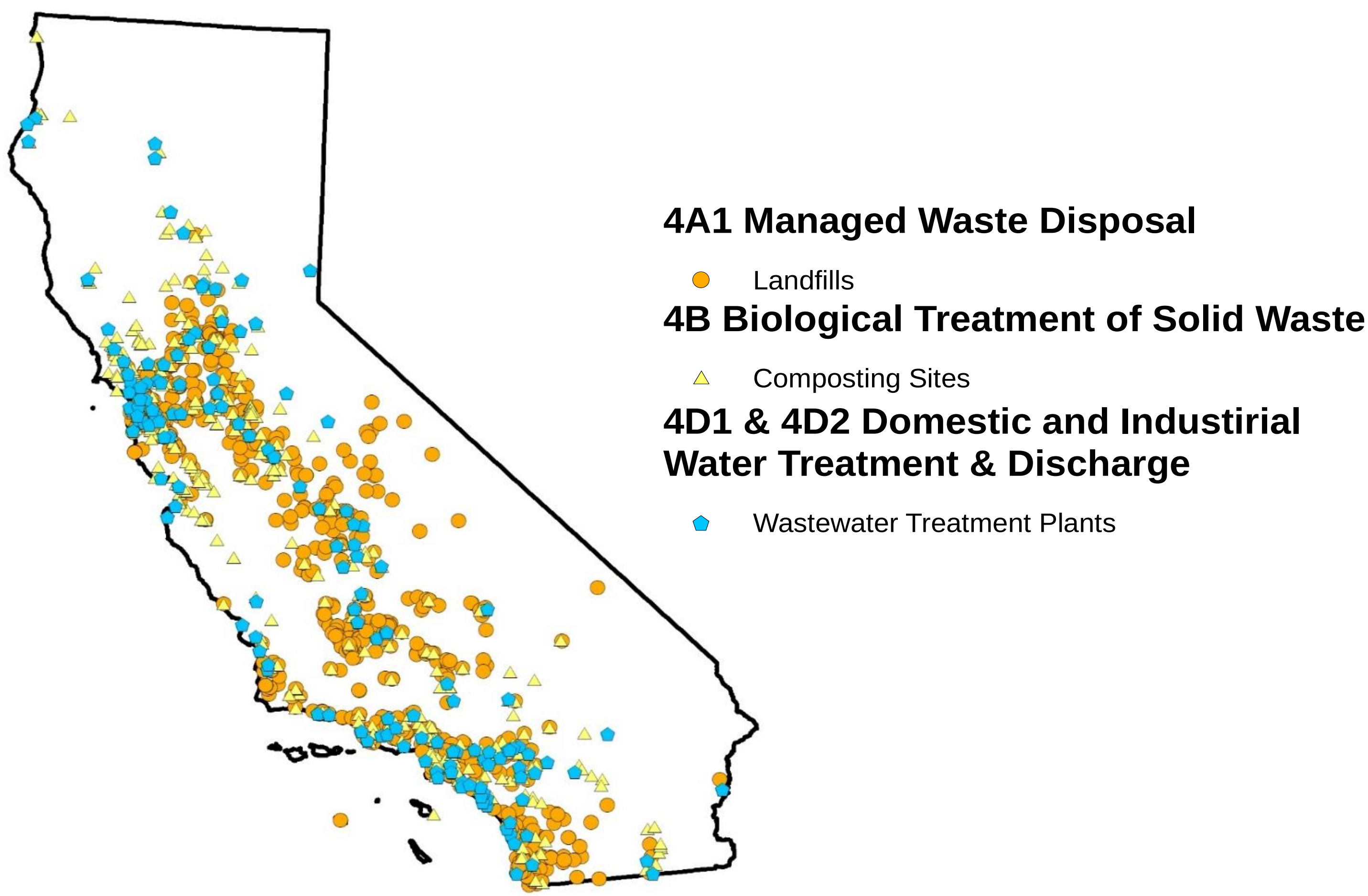


Figure 2.2 AVIRIS-NG Plumes - Vista-CA Unique Facilities Attribution. This map shows the locations of the 290 unique facilities as identified by the automated attribution of the 1,181 AVIRIS-NG plumes using Vista-CA GSAAM V4. The inset shows greater detail for the southern San Joaquin Valley and the Los Angeles Basin, where 1,043 plumes were observed.

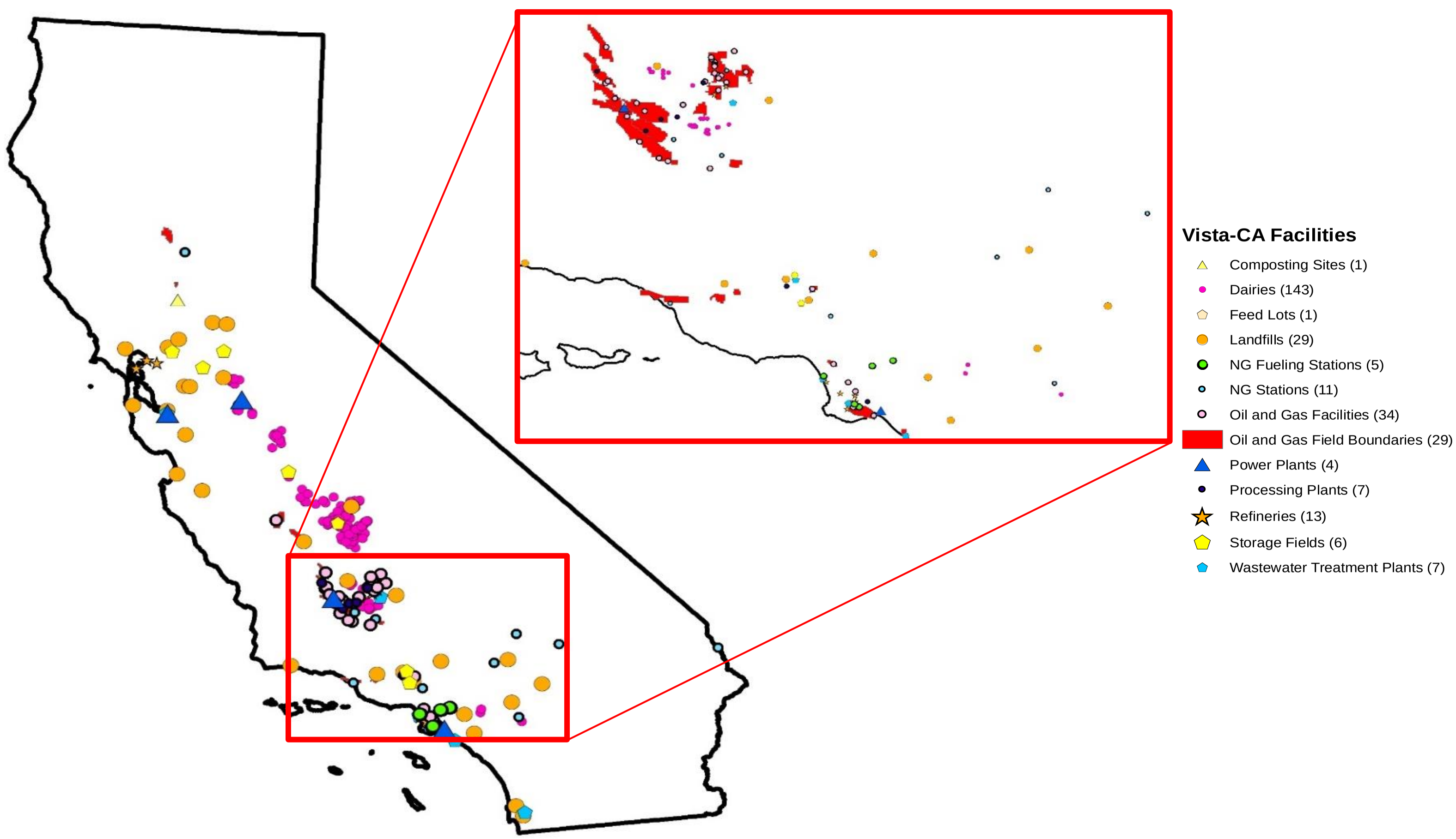


Table 2.1 Vista-CA Dataset. Vista-CA layers, representing CH_4 sources corresponding to IPCC Level 3, are shown organized by IPCC greenhouse gas emission reporting classification. Layers are grouped into facility and sub-facilities based on a systematic spatial overlap analysis and a logical process-based relationship. The vector formats and the search radii required for appropriate attribution for each layer are also given (See Chapter 2 Supplement).

IPCC Level 1	IPCC Level 3	Vista-CA Facility	Number of Elements	Vista-CA Sub-Facility	Number of Elements
1. ENERGY	1A1 Energy Industries	Refineries (polygons at 0m)	26	Sub-facility Oil & Gas Facility (polygons at 0m) Sub-facility Power Plants (polygons at 0m)	28 16
		Stand-Alone Power Plants (polygons at 0m)	326		
		sub-total	**352**	*sub-total*	**44**
	1B2 Oil & Natural Gas			Distribution Pipelines (polylines <11 m)	569,609
		NG Fueling Stations (polygons at <3 m)	209		
		NG Stations (points at 0m)	1,120		
		Oil & Gas Facilities (polygons at 0m)	3,261	Sub-facility Power Plants (polygons at 0m)	16
		Oil & Gas Fields (polygons at 0m)	463	Oil and Gas Wells (points at <25 m) Pipelines (polylines at <40 m)	211,735 (total 225,766) 15,149 (total 96,823)
		Processing Plants (polygons at 0m)	26	Sub-facility Oil & Gas Facility (polygons at 0m) Sub-facility Power Plants (polygons at 0m)	27 3
		Storage Fields (polygons at 0m)	12	Sub-facility Power Plants (polygons at 0m) Sub-facility Oil & Gas Facilities (polygons at 0m)	1 40
		sub-total	**5,091**	*sub-total*	**796,580**
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 Enteric Fermentation	Feed Lots (points at <1402 m)	72		
	3A1 & 3A2 Enteric Fermentation & Manure Management	Dairies (points at <1402 m)	1,715	Digesters (polygons <10 m) Sub-facility Power Plants (polygons)	31 2
		sub-total	**1,787**	*sub-total*	**33**
4. WASTE	4A1 Managed Waste Disposal	Landfills (polygons at <1 m)	714	Sub-facility Oil & Gas Fields (polygons at 0m) Sub-facility Power Plants (polygons at 0m) Sub-facility Wastewater Treatment Plant (polygons at 0m)	53 45 16
	4B Biological Treatment of Solid Waste	Composting Sites (polygons* at <230 m)	430		
		sub-total	**1,144**	*sub-total*	**114**
	4D1 & 4D2 Domestic and Industirial Water Treatment & Discharge	Wastewater Treatment Plants (polygons at 0m)	133	Power Plants (polygons at 0m)	26
TOTAL Vista-CA Facility Elements			**8,507**	**TOTAL Vista-CA Sub-Facility Elements**	**796,797**
TOTAL Vista-CA Elements			**901,009**		

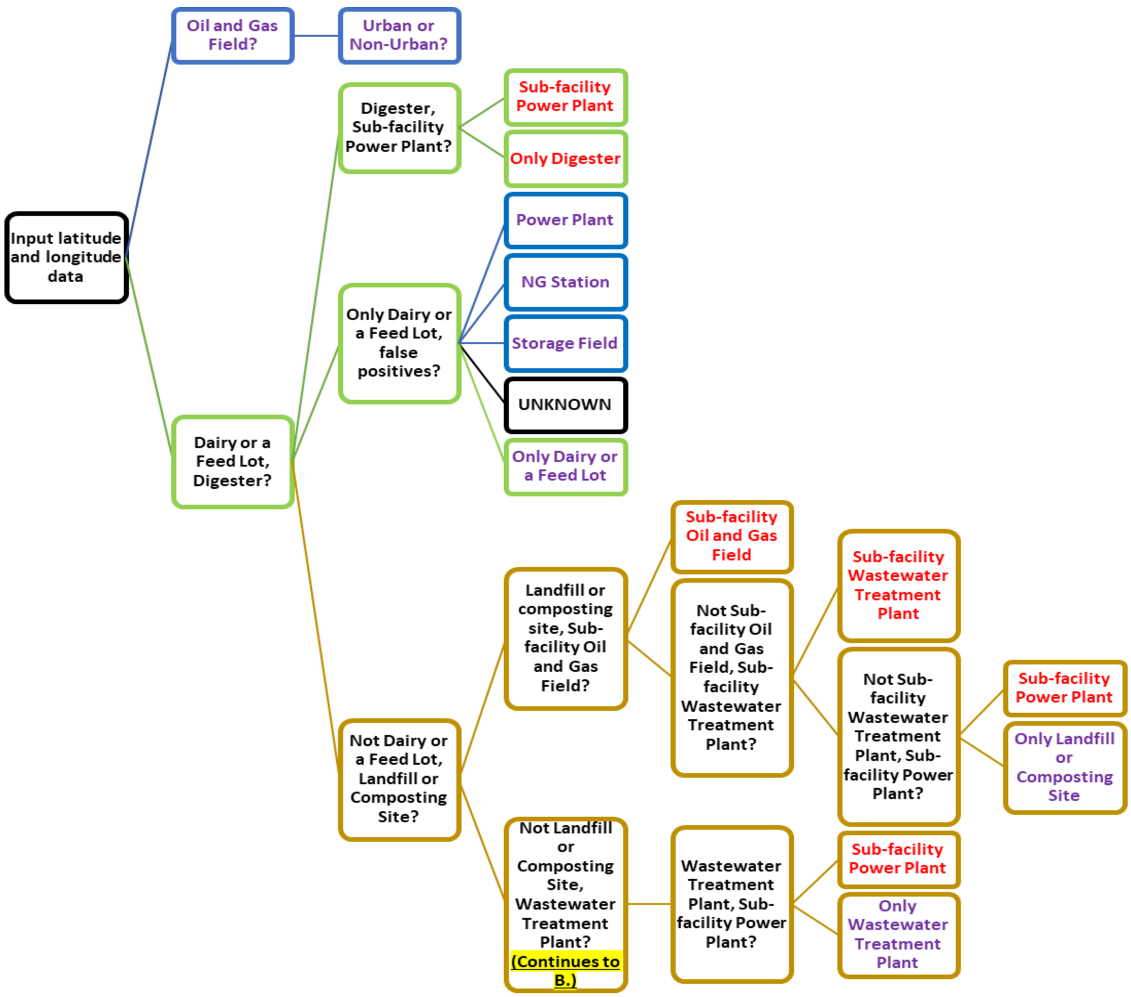


Figure 2.3 A.) Geospatial Source Attribution Automated Model (GSAAM) V4 Decision-Tree Logic. This figure illustrates the framework used to for the model logic. The model uses a spatial function at each node (Vista-CA dataset) and to attribute a methane plume observation to a source from amongst Vista-CA infrastructure elements based on a predetermined distance from the feature. Each pathway (attribution) ends with one of three options: Facility (purple text); Sub-facility (red text); or Unknown (black text). GSAAM is organized by IPCC categories: 1) Agriculture, Forestry & Other Land Use (green boxes); 2) Waste (orange boxes); and 3) Energy (blue boxes).

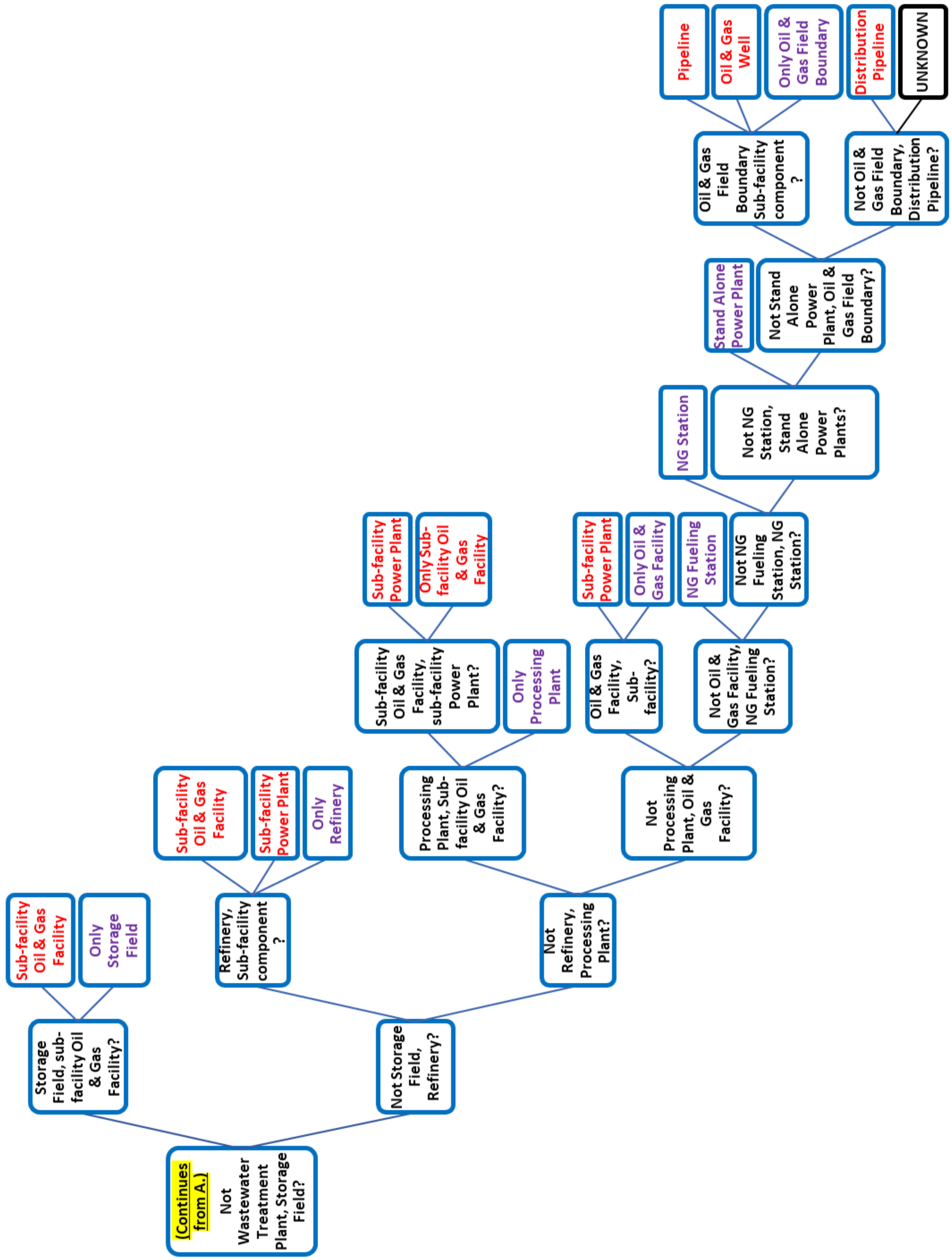

(Continues from A.) Not Wastewater Treatment Plant, Storage Field?
Storage Field, sub-facility Oil & Gas Facility?
Sub-facility Oil & Gas Facility
Only Storage Field
Not Storage Field, Refinery?
Refinery, Sub-facility component ?
Sub-facility Oil & Gas Facility
Sub-facility Power Plant
Only Refinery
Not Refinery, Processing Plant?
Processing Plant, Sub-facility Oil & Gas Facility?
Sub-facility Oil & Gas Facility, sub-facility Power Plant?
Sub-facility Power Plant
Only Sub-facility Oil & Gas Facility
Only Processing Plant
Not Processing Plant, Oil & Gas Facility?
Oil & Gas Facility, Sub-facility?
Sub-facility Power Plant
Only Oil & Gas Facility
Not Oil & Gas Facility, NG Fueling Station?
NG Fueling Station
Not NG Fueling Station, NG Station?
NG Station
Not NG Station, Stand Alone Power Plants?
Stand Alone Power Plant
Not Stand Alone Power Plant, Oil & Gas Field Boundary?
Oil & Gas Field Boundary Sub-facility component ?
Pipeline
Oil & Gas Well
Only Oil & Gas Field Boundary
Not Oil & Gas Field Boundary, Distribution Pipeline?
Distribution Pipeline
UNKNOWN

Figure 2.3 B.) Geospatial Source Attribution Automated Model (GSAAM) V4 Decision-Tree Logic. This figure illustrates the framework used to for the model logic. The model uses a spatial function at each node (Vista-CA dataset) and to attribute a methane plume observation to a source from amongst Vista-CA infrastructure elements based on a predetermined distance from the feature. Each pathway (attribution) ends with one of three options: Facility (purple text); Sub-facility (red text); or Unknown (black text). GSAAM is organized by IPCC categories: 1) Agriculture, Forestry & Other Land Use (green boxes); 2) Waste (orange boxes); and 3) Energy (blue boxes).

Table 2.2 Vista-CA GSAAM-AVIRIS-NG Attribution Results. Maximum Source Attribution Potential is calculated by taking the total potential plumes that have been attributed with Vista-CA and adding the total number of false negatives. The facility and sub-facility accuracies are calculated by taking the total correct attributions for each and subtracting the number of unknown attributions. Unknown attributions show the total number of plume attributions that were unable to be identified by GSAAM using Vista-CA datasets. The False Positive Rate is measured by taking the total number of plumes falsely attributed to another Vista-CA feature. The False Negative Rate is measured by taking the total number of plumes that were not attributed to a Vista-CA facility even though a Vista-CA facility for them logically exists.

Data Source	Maximum Source Attribution Potential		Vista-CA Facility Accuracy		Vista-CA Sub-Facility Accuracy		Number of Unknown Attributions		False Positive Rate		False Negative Rate	
AVIRIS-NG Unpublished Fall 2016-Fall 2017 Campaign (748 Plumes)	744/748	99.47%	738/740	99.73%	118/118	100%	8/748	1.07%	2/748	0.27%	4/748	0.53%
AVIRIS-NG Published Fall 2016-2017 Campaign (1,181 Plumes)	1170/1,181	99.07%	1158/1,170	98.97%	271/271	100%	11/1,181	0.93%	1/1181	0.08%	0/1181	0.00%
AVIRIS-NG Unpublished Fall 2018 Campaign (495 Plumes)	493/495	99.59%	488/493	98.99%	79/79	100%	2/495	0.40%	5/495	1.01%	0/495	0.00%
OVERALL PERFORMANCE	**99.38%**		**99.23%**		**100%**		**0.87%**		**0.45%**		**0.18%**	

IPCC Level 1	IPCC Level 3	Vista-CA Categories	CARB PMT	EPA FLIGHT	Vista-CA
			Number of Facilities/Features	Number of Facilities/Features	Number of Facilities/Features
1. ENERGY	**1A1 Energy Industries**	Refineries	24 (13)	19 (8)	26 (13)
		Power Plants (Facility & Sub-facility)	209 (7)	113 (5)	433 (4)
		sub-total	***233 (20)***	***132 (13)***	***459 (17)***
	1B2 Oil & Natural Gas	Distribution Pipelines (Sub-facility)	0 (N/A)	3 (N/A)	569,609 (N/A)
		NG Fueling Stations	0 (0)	0 (0)	209 (5)
		NG Stations	13 (6)	11 (3)	1,120 (11)
		Oil & Gas Facilities	39 (8)	0 (0)	3,356 (34)
		Oil & Gas Fields	0 (0)	0 (0)	516 (29)
		Oil & Gas Wells (Sub-facility)	0 (N/A)	0 (N/A)	225,766 (N/A)
		Pipelines (Sub-facility)	10 (N/A)	3 (N/A)	96,823 (N/A)
		Processing Plants	0 (0)	7 (1)	26 (7)
		Storage Fields	0	3 (2)	12 (6)
		sub-total	***62 (14)***	***27 (6)***	***897437 (92)***
		sub-total	***295 (34)***	***159 (19)***	***897,896 (109)***
3. AGRICULTURE, FORESTRY & OTHER LAND USE	**3A1 Enteric Fermentation**	Feed Lots	0 (0)	0 (0)	72 (1)
	3A1 & 3A2 Enteric Fermentation & Manure Management	Dairies	0 (0)	0 (0)	1,715 (143)
	3A2 Manure Management	Digesters (Sub-facility)	0 (N/A)	0 (0)	33 (N/A)
		sub-total	***0 (0)***	***0 (0)***	***1,320 (144)***
4. WASTE	**4A1 Managed Waste Disposal**	Landfills	14 (13)	99 (24)	714 (29)
	4B Biological Treatment of Solid Waste	Composting Sites	0 (0)	0 (0)	430 (1)
	4D1 & 4D2 Domestic and Industirial Water Treatment & Discharge	Wastewater Treatment Plants	14 (5)	1 (1)	149 (7)
		sub-total	***28 (18)***	***100 (25)***	***1,293 (37)***
TOTAL Number of Facilities/Features			**331 (52)**	**263 (44)**	**901,009 (290)**

Table 2.3 Facilities/Features Comparison. Number of facilities/features in CARB PMT, EPA FLIGHT, and Vista-CA have been categorized and binned into specific Vista-CA categories for comparison. The number of unique facilities for which CH_4 observations by AVIRIS-NG have been attributed for each dataset is shown in parenthesis.

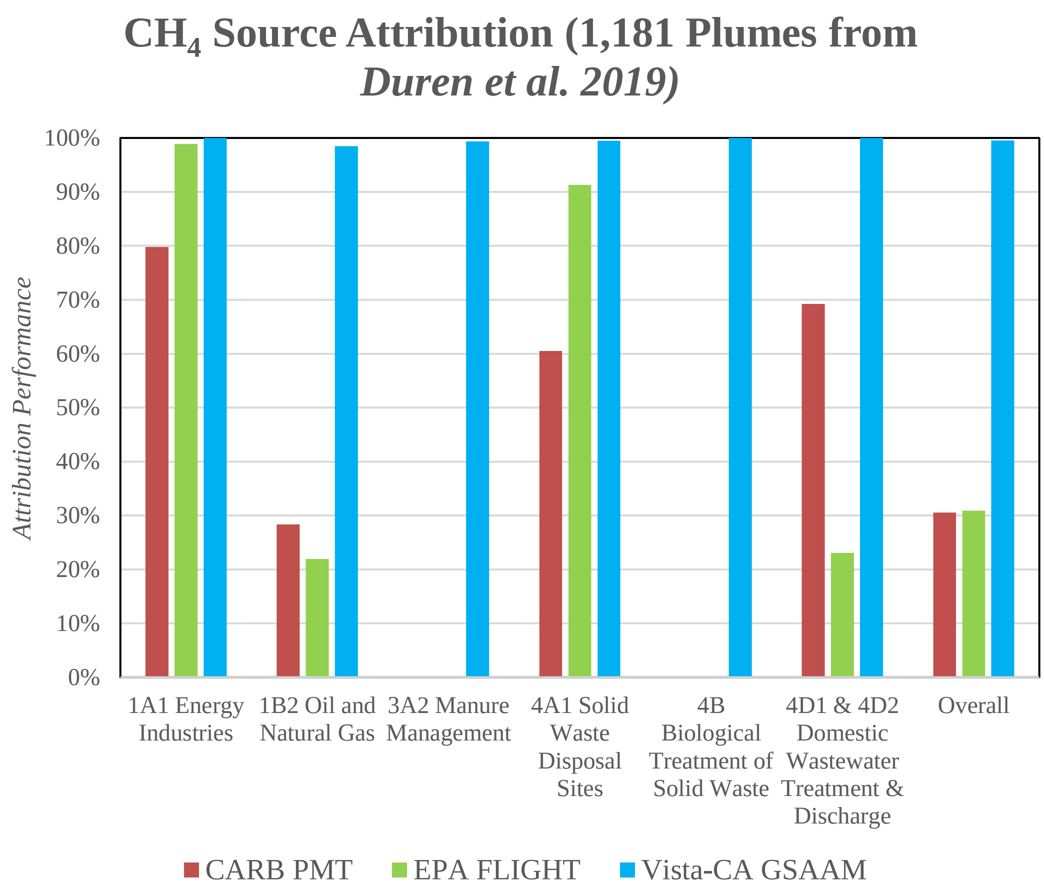


Figure 2.4 Overall Attribution Accuracy Results. Manual attribution performance as a percentage for CARB PMT, EPA FLIGHT, and Vista-CA GSAAM attribution model performance for IPCC Level 3 source categories. (3A1 Enteric Fermentation: 0 Attributions detected for Feed Lots).

V. Supplementary

S1. Introduction

This document outlines the specific steps and procedures conducted to develop and construct the data, methods, and algorithms summarized in the main paper. The datasets and algorithm are accessible at https://doi.org/10.3334/ORNLDAAC/1726. Intergovernmental Panel on Climate Change (IPCC) and California Air Resources Board (CARB) inventories are organized by commercial and industrial economic activity that results in greenhouse gas (GHG) emissions. IPCC Level 1 base categories: (1) energy, (2) agriculture, and (3) waste, broadly represent different emission source sectors. For Vista-CA the primary level of organization are spatial layers comprising groups of individual facilities that are fixed locations of identified infrastructure or an area where the activities of processing and production occur organized by source type, and then categorized by sector. Facilities can contain many sub-facility system infrastructure and components. The sub-facility category refers to identified infrastructure components which are located within a facility-system that directly or indirectly aid in operation and can spatially overlap with the identified facility. This categorization scheme accounts for duplicate sub-facility infrastructure components occurring within different facilities across different sectors. This setup also allows for the start of process-based understanding of each sector for emissions prediction and quantification. For example, a dairy digester can be considered a sub-facility of a dairy, or a power plant can be considered a sub-facility level feature for a larger refinery or wastewater treatment plant.

S2. Vista-CA Datasets

All Vista-CA datasets underwent significant QA/QC using a variety of GIS methods (Figure S2.1). We utilized ArcGIS 10.6.1 to conduct all GIS related tasks including dataset development. All datasets are available in ESRI shapefile format. Google Earth and ESRI Basemaps were the primary aerial image sources used for validation and polygon generation. All Vista-CA datasets contain Vista-related metadata which include Vista-ID, Vista Date, Vista IPCC, Vista Name, Vista Source, and Vista Source Type. Vista-ID is a unique alphanumeric feature identifier for every feature in the Vista database. Vista Date contains the most recent date of data update and change. Vista IPCC identifies the IPCC sector designation of a given Vista feature. Vista Name is the name of the given Vista facility or infrastructure. Vista Source outlines the original source that the respective Vista dataset was developed from and Vista Source Type identifies the infrastructure/sector type for each Vista-CA dataset. The following sections provide a detailed outlined of major updates and changes made to specific Vista-CA datasets along with the procedures undertaken to perform those tasks.

S2.1 Vista-CA Composting Sites

Vista-CA Composting Sites were developed as a subset from the 2015 CalRecycle data. We decided on incorporating these landfills types because of observed emissions by AVIRIS-NG. Based on activity categories, we filtered and maintained the following Composting Site types: Biosolids Composting at Publicly Owned Treatment Works (POTWs), Chipping and Grinding Activity Facility/Operation, Composting Facility (Green Waste), Composting Facility (Mixed), Composting Facility (Other), Composting Facility

(Sludge), Composting Operation (Ag), Composting Operation (Green Waste), and Composting Operation (Research). In total there were 430 facilities that were included in the final Vista-CA Composting Site dataset.

S2.2 Vista-CA Dairies

Vista-CA Dairies and Vista-CA Feed Lots were developed by matching physical locations of dairy farms with information obtained from permits granted by several government agencies that regulate water and air emissions from dairies and contextualized through aerial imagery. We obtained annual reports from 909 (420 facilities from Rancho Cordova and 489 facilities from Fresno) dairies from the California Regional Water Quality Control Board (RWQCB) Central Valley (Region 5) and Santa Ana (Region 8) offices for 2015 (Carranza et al., 2018). These reports contain the name, physical address, coordinates, herd size, and management information related to the nutrient budget of dairy farms located in the Central Valley. State regulations require all dairies in this region to submit this data annually for purposes of regulating waste discharge that could impact water quality (CALIFORNIA REGIONAL WATER QUALITY CONTROL BOARD CENTRAL VALLEY REGION Order No. R5-2007-0035 WASTE DISCHARGE REQUIREMENTS GENERAL ORDER FOR EXISTING MILK COW DAIRIES: https://www.waterboards.ca.gov/centralvalley/board_decisions/adopted_orders/general_orders/r5-2007-0035.pdf). These regions account for more than 95% of milk production in the state (CDFA, 2017), and include the counties of Butte, Fresno, Glenn, Kern, Kings, Madera, Merced, Riverside, Sacramento, San Bernardino, San Joaquin, Siskiyou, Solano, Stanislaus, Tehama, Tulare, Yolo, Yuba, and others containing no dairies (for a complete list, see https://www.waterboards.ca.gov/publications_forms/publications/factsheets/docs/region_brds.pdf).

We also used air permit data collected by the San Joaquin Valley Air Pollution Control District (SJAPCD) under California Senate Bill 700, 2003 (http://www.leginfo.ca.gov/pub/0304/ bill/sen/ sb_0651-0700/sb_700_bill_20030922_chaptered.html). These permits are required for dairies with more than ~1954 cows (https://www.valleyair.org/farmpermits/updates/draft_dairy_bact.pdf). These permits contain information about the location of the facility including addresses and coordinates, maximum number of cows the facility could house, and manure management infrastructure. This data included information for 1,036 dairies, 69% (718/1,036) of which overlapped with the RWQCB permit facilities.

Finally, we obtained a list of permitted facilities from the California Integrated Water Quality System (CIWQS) Project (e.g., National Pollutant Discharge Elimination System surface water discharge permits required under the Clean Water Act). CIWQS provides the facility name, addresses, and coordinates, for all active, permitted dairies in the state (CIWQS Regulated Facility Reports, https://ciwqs.waterboards.ca.gov/ciwqs/readOnly/CiwqsReportServlet?in Command=reset&reportName=RegulatedFacility). We used CIWQS records of facility name and address for 330 dairies located outside of the central valley. CIWQS R5 included 1,520 dairies across the state. Both CIWQS and CIWQS datasets contained information on coordinates, addresses, and names. Additionally, 56% (844/1,520) of CIWQS R5 dairies overlapped with both RWQCB and SJAPCD dairies.

We determined the locations of dairy farms in California from satellite imagery viewed primarily in Google Earth. Secondary sources include Google Maps, ESRI Basemaps and tertiary sources were through sites such as Manta.com that included facility/business records online. High-resolution aerial imagery permits geolocation of dairy farms due to their distinctive appearance, typically with long white roofs for shelter, dark colored square or rectangle manure lagoons, and staggered corral areas located together in proximity. These spatial features enabled identifications of facilities. Using this methodology, we manually isolated and identified point locations for 1,709 dairy-related facilities across the state (RAFIQ). RAFIQ data only includes geographic coordinates and reverse geocoded address information. We utilized the following hierarchy to prioritize inclusion and validation from the data sources described above: RWQCB > SJAPCD > CIWQS > CIWQS R5 > RAFIQ. All data from RWQCB was maintained, while only unique (non-overlapping) data were maintained from subsequent datasets in the hierarchy. First, data from all 909 RWQCB dairies were examined and duplicate entries of facility names and facility addresses along with missing or incorrect spatial data were identified and removed resulting in 850 total remaining dairies maintained from this dataset. Next, data from 318 unique SJAPCD dairies were examined and duplicate entries of facility names and facility addresses along with missing or incorrect spatial data were identified and removed resulting in 274 total remaining dairies stored from this dataset. Data from 330 unique CIWQS dairies were examined and duplicate entries of facility names and facility addresses along with missing or incorrect spatial data were identified and removed resulting in 311 total remaining dairies maintained from this dataset. Similarly, data from 676 unique CIWQS R5 dairies were examined and duplicate entries of facility names and facility addresses along with missing or incorrect spatial data were identified and removed resulting in 218 total remaining dairies stored from this dataset. Lastly, the remaining dataset resulted in a total of 1,653 dairies (RWQCB, SJAPCD, CIWQS, CIWQS R5, RAFIQ). These 1,653 dairies spatially overlapped against 91.3% (1,562/1,709) of RAFIQ dairies and the remaining 127 RAFIQ dairies were appended to the RSCC dataset. In order to constrain the domain or spatial extent of each facility, dairies were geocoded using their coordinates and addresses in order to determine the approximate location of each dairy. Using contextual aerial imagery, point vectors were manually adjusted to reflect the center of a given dairy facility (typically overlaid on feed lot features) and coordinates were recalculated and updated for the entire dataset. The final dairy facility dataset contains 1,780-point locations across California. A unique 9-digit alphanumeric identifier, the Vista-identification number (Vista-ID), was attributed to all 1,780 facilities for data management and organizational purposes.

A query was used to extract relevant feed lot designated facilities from the point locations of 1,780 dairies. Facilities with names that contained any of the following phrases or words were extracted: “beef”(2), “cattle”(29), “cattle ranch”(1), "feeders”(2), "feed lot"(2), "feedlot"(5), "feed yard"(1), "feedyard”(3), “livestock”(7), “livestock market”(3), and “stockyard”(1), other (3). This resulted in an extraction of 59 facilities. The Vista-ID of the 58 extracted feed lot facilities was exported and subsequently joined to the 1,780 dairy dataset based on the Vista-ID in order to maintain geospatial and attribute records during final extraction. Matched Vista-ID records were extracted and stored as a separate dataset for the 72 facilities in the Vista-CA Feed Lot layer. Remaining records of 1,708 dairies in the final dairy dataset were extracted from the 1,780 point dataset as an independent portion for the Vista-CA Dairy layer and 7 additional dairies were added to the dataset for a total of 1,715 dairies. Every facility in both final Vista datasets contains appended Vista metadata which includes a Vista-ID, source of the data, IPCC sector designation (3A1 Enteric Fermentation for Feed Lots and 3A1 & 3A2 Enteric Fermentation and Manure Management), source type description, and the last date of update.

S2.3 Vista-CA Digesters

Vista-CA Digester data was collected from the 2016 Environmental Protection Agency's database on Agricultural Biogas Recovery program called AGSTAR. This dataset contained 27 point locations of digesters in California of which 21 were validated using aerial imagery. Digesters located on 12 farms were further found using aerial imagery and were added to the dataset. All points were converted to polygons based on the physical locations of relevant digester infrastructure as seen using aerial imagery. In total there were 33 digesters validated with associated polygons generated for the Vista-CA Digester dataset.

S2.4 Vista-CA Landfills

Vista-CA Landfill data was collected from the 2018 Vista-LA published dataset (Carranza et al. 2018), the 2015 CalRecycle data, the 2015 CARB dataset along with the 2016 EPA FLIGHT dataset. Out of the original 3,087 landfills we selected the following activity categories: Asbestos Containing Waste (ACW) Site, Construction and Demolition (CDI) Waste Disposal Facility, Engineered Municipal Solid Waste (EMSW) Conversion, Industrial Waste Codisposal Facility, Inert Debris ENG Fill Operation, Inert Waste Disposal Site, Land Application, Limited Volume In-Vessel Digestion, Medium Volume In-Vessel Digestion, Solid Waste Disposal Site, Solid Waste Landfill, Treatment Unit (in situ) resulting in a total of 714 landfills. Each landfill was then validated and a polygon extent was created using aerial imagery. Based on AVIRIS-NG retrievals we created placeholder polygons for 3 landfills located in isolated/rural oil fields (Buena Vista Hills Disposal Site LNF000083, Arco Fairfield Lease Disposal Facility LNF000030/Derby Acres Bd LNF000200 (overlapped), and Kern Front Disposal Site LNF000335).

S2.5 Vista-CA Natural Gas Fueling Stations

Vista-CA Natural Gas Fueling Station data was obtained from the 2017 U.S. Department of Energy’s Alternative Fuels Datacenter portal. This dataset contained information on 162 compressed natural gas fueling stations and 46 liquefied natural gas fueling stations. We identified the pertinent infrastructure using high-resolution aerial imagery and generated polygons around them. Finally, we combined all 209 stations from both datasets into the Vista-CA Natural Gas Fueling Station layer.

S2.6 Vista-CA Natural Gas Stations

New data for 1,120 natural gas compressor stations was collected from the California Energy Commission (CEC) along with 10 stations from the EPA FLIGHT tool. Each individual station was validated using aerial imagery and standardized with Vista metadata. Attempted to create polygons for points but resulted in mediocre success, and reverted back to validating only points.

S2.7 Vista-CA Oil & Gas Field Boundaries and Facilities

Oil and gas facility and field boundaries, not previously in Vista-LA, were additionally added to the database as well. Data was obtained from the Division of Oil, Gas, and Geothermal Resources for the entire state for 2018. There were 516 field boundaries and 3,356 facility boundaries in the datasets. Each dataset was georeferenced and standardized with Vista metadata.

S2.8 Vista-CA Oil & Gas Wells

Vista-CA Oil and Gas Well data was obtained from the Division of Oil, Gas, and Geothermal Resources in 2017 (DOGGR). This dataset contained 226,652 wells. We identified 886 wells that contained insufficient geolocation metadata information which we omitted, leaving a total of 225,766 wells for the Vista-CA Oil and Gas Well dataset.

S2.9 Vista-CA Pipelines & Distribution Pipelines

Data for natural gas pipelines includes information on both transmission and distribution networks. The natural gas pipeline dataset was derived by appending unique lines from the Pipeline and Hazardous Materials Safety Administration (PHMSA) 2013 dataset, 2012 California Energy Commission (CEC) dataset, and the 2016 U.S. Energy Information Administration (EIA) dataset. A separate urban residential distribution line network was geospatially constructed using parts of the 2018 California road network (U.S Census Bureau) and was overlaid on the 2018 NLCD landcover dataset. Raster cells that were classified as being 20-100% impervious in urban areas were extracted and used to define the extent of the road network. The resulting lines were then spatially connected to the existing California natural gas pipeline infrastructure network using a 10km proximity tolerance. Vista metadata was appended for each line segment in the network. In total, there were 216,774 km of lines for the entire network of which 196,670 km were derived from the California road/NLCD network and 20,104 km were from the NPMS/CEC/EIA network derived dataset.

S2.10 Vista-CA Wastewater Treatment Plants

Vista-CA wastewater treatment plant data was collected from the 2018 Vista-LA dataset and the 2016 CARB dataset. Combined, there were 148 point locations for wastewater facilities. We then converted all points to polygons through aerial image assessment. Using AVIRIS-NG observations we identified one additional facility and added it to our dataset. In total, the Vista-CA Wastewater Treatment Plant dataset contains 149 validated polygons.

S3. Geospatial Source Attribution Automated Model (GSAAM)

The geospatial source attribution model automatically and efficiently attributes Vista-CA potential methane emitting facilities and infrastructures to latitude/longitude point source retrievals. Built on ESRI ArcGIS Model Builder 10.6.1 and Python 2.7.14, this model automates the identification of a methane plume through proximity-based spatial analysis using a systematic decision-tree scheme incorporating Vista-CA data.

Vista-CA sectors were implemented in the decision-tree based on overall spatial patterns and relation to plume occurrence persistence. Vista-CA facility ID’s are logically assigned to each latitude/longitude plume occurrence. Currently this automated model yields an accuracy measure of 99% for attributing a Vista-CA facility to a methane plume occurrence. Accuracy measure was based on the number of matches found between the automated source attribution result and the AVIRIS-NG manual source identification effort.

Spatial biases taken into account for GSAAM development included spatial density, spatial extents, and spatial overlaps (Figures S2.2). Spatial density refers to the number of features in a given area. GSAAM accounts for sub-facilities with high spatial density such as Oil and Gas Wells and facilities like Natural Gas Storage Fields that are low spatial density cases. Spatially dense Vista-CA datasets can potentially yield false positives (attribution result which incorrectly indicates that a particular plume is attributed to a sector, facility, or sub-facility) while low-density sectors can lead to false-negatives (attribution result which incorrectly misses an attribution for a particular plume to a sector, facility, or sub-facility). GSAAM accounts for spatial extent which evaluates the amount of geographic space one Vista-CA feature occupies or covers compared to a different feature. An example of a feature with a large spatial extent would be the Oil and Gas Field Boundaries sector which covers a footprint of tens of hundreds of square kilometers, while smaller spatial extents would be exemplified in the Oil and Gas Wells covering about a square meter at most on the ground. A larger spatial extent has the potential to lead to more false positive cases, while a smaller spatial extent feature can potentially result in false negatives. GSAAM Version 4 accounts for spatial extent biases by individually incorporating a field for Oil and Gas Field Boundaries as we recognize the spatial biases associated with coverage (Figure S2.2). Finally, an extensive spatial overlap analysis was conducted to spatially recognize any potential facility and sub-facility associations among the Vista-CA datasets (Table S2.2).

The basic geospatial functions used in GSAAM include Select, Near, Join Field, Merge, Feature to Point, Make XY Event Layer, Delete Field, Add Field, Calculate Field, and Table to Excel of which the first 4 are the most important as it relates to model function. Select extracts features from the input feature using Structured Query Language (SQL) expression and stores them in the output. This is essentially used to partition the plume data for attribution to the Vista-CA datasets based on the predetermined hierarchy (Figure 2.3). Near calculates distance and additional proximity information between the input and the closest feature in another layer. The near function is the crux of the proximity attribution framework as it relates plumes to the nearest Vista-CA layer (Figure S2.2). Join Field joins the contents of one table in a feature to another table based on a common attribute field, which is used to associate Vista-ID's to individual plumes. Merge combines multiple input datasets into a single, new output dataset, which is used to combine all the attributions and non-attributions back into one table as the final result. As a simple example, for one given attribution assessment against one Vista-CA layer, this is the order of the model workflow: Make XY Events Layer > Feature to Point > Delete Field > Add Field > Calculate Field > Near > Select > Join Field > Merge > Table to Excel. GSAAM is an ArcGIS Executable Function that can either be run using ArcGIS or independently in a Python console/shell.

The end result of GSAAM produces a Microsoft Excel comma separated value (.csv) file that contains 6 new fields (in addition to any existing fields including latitude and longitude) (Table S2.1): IPCC Level 1, IPCC Level 3, Vista-CA Facility, Vista-CA Sub-Facility, Oil and Gas Field, and Urban OG Field. IPCC Level 1 and 3 are the first and third levels of the IPCC framework that Vista-CA is based off of. Vista-CA Facility and Vista-CA Sub-Facility fields respectively provide the Vista-ID associated to that individual plume as determined by the algorithm. Oil and Gas Field identifies the Vista-ID for the Oil and Gas Field that is spatially associated with any plume points. Finally, Urban OG Field column determines whether or not the Oil and Gas Field associated to a plume falls in an urban area.

Figure S2.1 Vista-CA Dataset Summary. Includes the source and timeframe of the original dataset, the maximum spatial coverage of each original dataset, the number of features/infrastructures associated with methane emissions of each Vista-CA layer, and the final Vista deliverable format they are each in.

*Sources may also include fugitive emissions that fall under IPCC source type 1B

[a] Source accessible only through correspondence and approval with publishers/authors

[b] Source not currently included in the California Air Resources Board's 2010-2017 GHG Inventory

IPCC Level 1	IPCC Level 2	Vista-CA Layers (IPCC Level 3)	Data Source (Year)	Original Data Spatial Coverage	Vista-CA No. of Features	Vista-CA Data Format
1. Energy	1A Fuel Comustion Activities	**Energy Industries (IPCC - 1A1)**				
		Power Plants	EIA (2016), SCAG (2005, 2012)	CONUS California	433	shapefile polygons
		Refineries	EIA (2016), SCAG (2005, 2012)	CONUS California	26	shapefile polygons
	1B Fugitive Emissions From Fuels**	**Oil and Natural Gas (IPCC - 1B2)**				
		Compressor Stations	CEC (2017) EPA FLIGHT Tool (2016)	California CONUS	1,120	shapefile polygons
		Distribution Pipelines	NLCD (2018) U.S. Census (2018)	California CONUS	569,609	shapefile polylines
		Natural Gas Fueling Stations*	U.S. DOE AFDC (2017)	CONUS	208	shapefile polygons
		Oil and Gas Facilities	CEC (2018)	California	3,356	shapefile polygons
		Oil and Gas Field Boundaries	CEC (2018)	California	516	shapefile polygons
		Oil and Gas Wells	DOGGR (2017)	California	225,766	shapefile points
		Pipelines	CEC (2012) NPMS (2013)	California CONUS	96,823	shapefile polylines
		Processing Plants	EIA (2014)	CONUS	26	shapefile polygons
		Storage Fields	DOGGR (2016) EIA (2016)	California CONUS	12	shapefile polygons
3. Agriculture, Forestry & Other Land Use	3A Livestock	**Enteric Fermentation (IPCC - 3A1)**				
		Feed Lots	CIWQS (2018) RWQCB (2017) SJAPCD (2017)	California Rancho Cordova and Fresno Regions San Joaquin Valley	72	shapefile points
		Manure Management (IPCC - 3A2)				
		Digesters	CARB (2016)	California	33	shapefile polygons
		Enteric Fermentation and Manure Management (IPCC - 3A1 and 3A2)				
		Dairies	CIWQS (2018) RWQCB (2017) SJAPCD (2017)	California Rancho Cordova and Fresno Regions San Joaquin Valley	1,715	shapefile points
4. Waste	4A Solid Waste Disposal	**Managed Waste Disposal (IPCC - 4A1)**				
		Landfills	CalRecycle (2015) SCAG (2005, 2012)	California California	714	shapefile polygons
	4B Biological Treatment of Solid Waste	**Biological Treatment of Solid Waste (IPCC - 4B)**				
		Composting Sites	CalRecycle (2015)	California	430	shapefile points
	4D Wastewater Treatment & Dishcarge	**Domestic and Industrial Water Treatment & Discharge (IPCC - 4D1 and 4D2)**				
		Wastewater Treatment Plants	CARB (2016) SCAG (2005, 2012)	California California	149	shapefile polygons

NOTE:
CalRecycle = California Department of Resources Recycling and Recovery
CARB = California Air Resources Board
CEC = California Energy Commission
CONUS = Contiguous United States Region
DOE AFDC = U.S. Department of Energy Alternative Fuels Data Center
DOGGR = California Department of Conservation, Division of Oil, Gas, and Geothermal Resources
EIA = U.S. Energy Information Administration
EPA FLIGHT Tool = U.S. Environmental Protection Agency Facility Level GreenHouse gas Tool
NLCD = U.S. Geological Survey National Land Cover Database
NPMS = National Pipeline Mapping System
RWQCB = California EPA Regional Water Quality Control Board, Fresno, Rancho Cordova, and Santa Ana Region
SCAG = Southern California Association of Governments
SJAPCD = San Joaquin Valley Air Pollution Control District

Figure S2.2 A) Aspects of Spatial Attribution: False Positive and False Negatives. The following biases were accounted for to reduce uncertainty in the attribution of CH_4 plumes. False Positives and False Negatives: This false positive example (left) illustrates how a dairy farm (yellow point) is adjacent to a plume detected on separate farm (green point), yet the attribution algorithm assigned the source of that plume to be the dairy farm, which would lead to an inaccurate attribution. This false negative example (right) shows that a plume was located outside of the immediate boundary of an Oil and Gas Field (tan polygon) and based on the preexisting attribution logic the plume (green point) was labeled as “Unknown” despite being adjacent to the Vista-CA Oil and Gas Field Boundary feature in an isolated oil and gas production region.

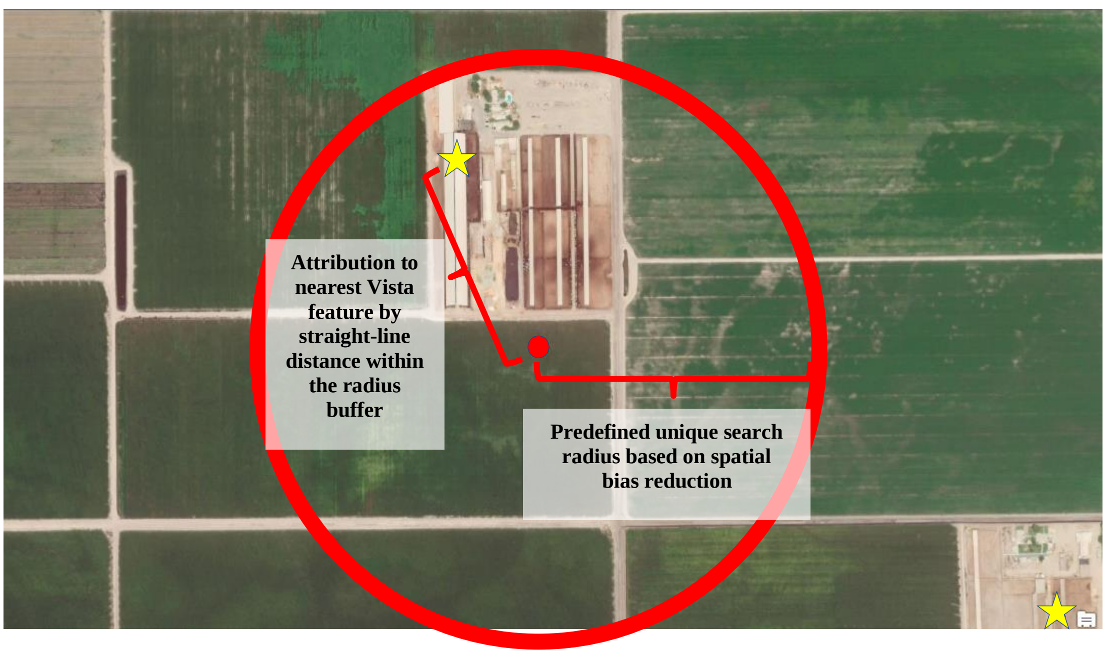


Figure S2.2 B) Aspects of Spatial Attribution: Proximity Spatial Attribution: The following biases were accounted for to reduce uncertainty in the attribution of CH_4 plumes. Using a predefined radius (red circle) from the plume point (red point) based on the relevant CH_4 emitting sector in order to search for the nearest Vista-CA feature (yellow star) within that given radius for the purposes of logical spatial attribution.

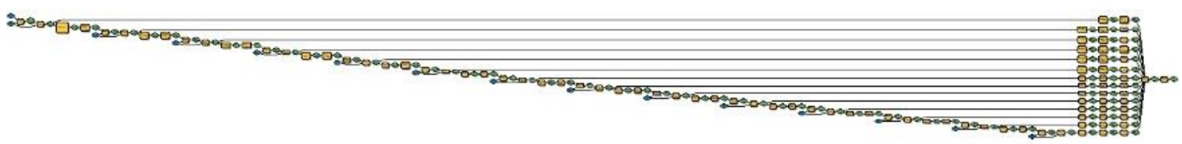


Figure S2.3 A) Geospatial Source Attribution Automated Model (GSAAM) Versions 1: Simple decision-tree model that contained a hierarchy based on limiting false-positives and preventing false-negatives but lacked inclusion and distinction of facility and sub-facility attributions.

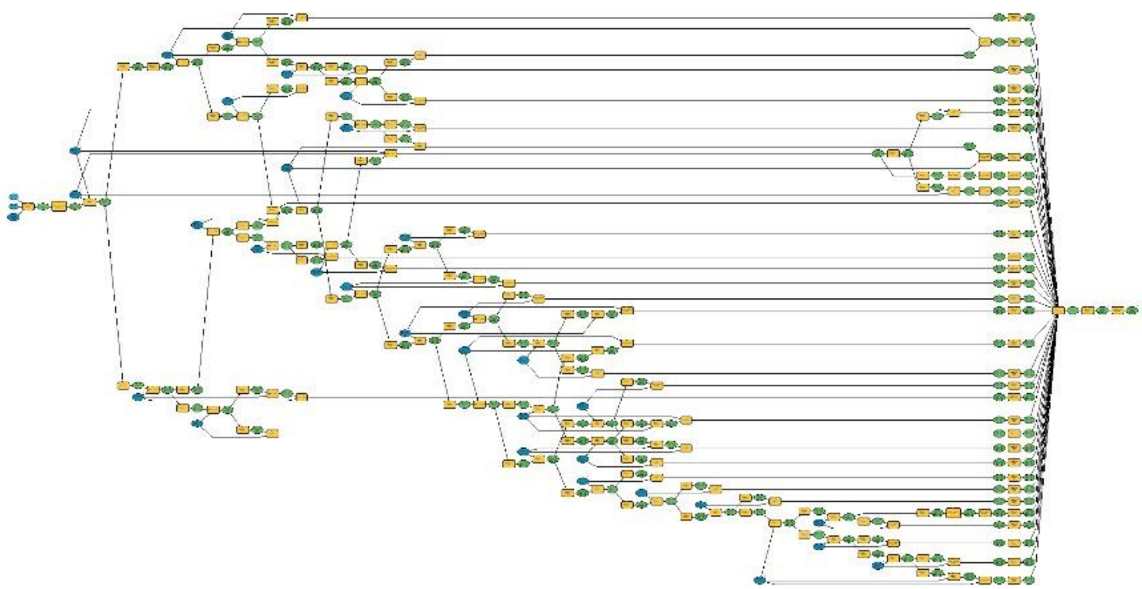


Figure S2.3 B) GSAAM Version 2: Multiple decision-tree model that restructured the Version 1 hierarchy by systematically incorporating spatial overlap and included sub-facility attribution but lacked distinction between sub-facility and facility attributions.

Figure S2.3 C) GSAAM Version 3: Multiple decision-tree model that incorporated all aspects of Version 2 and added functionality to differentiate between facility and sub-facility attributions while also allowing for IPCC Level 1 and Level 3 attribution outputs.

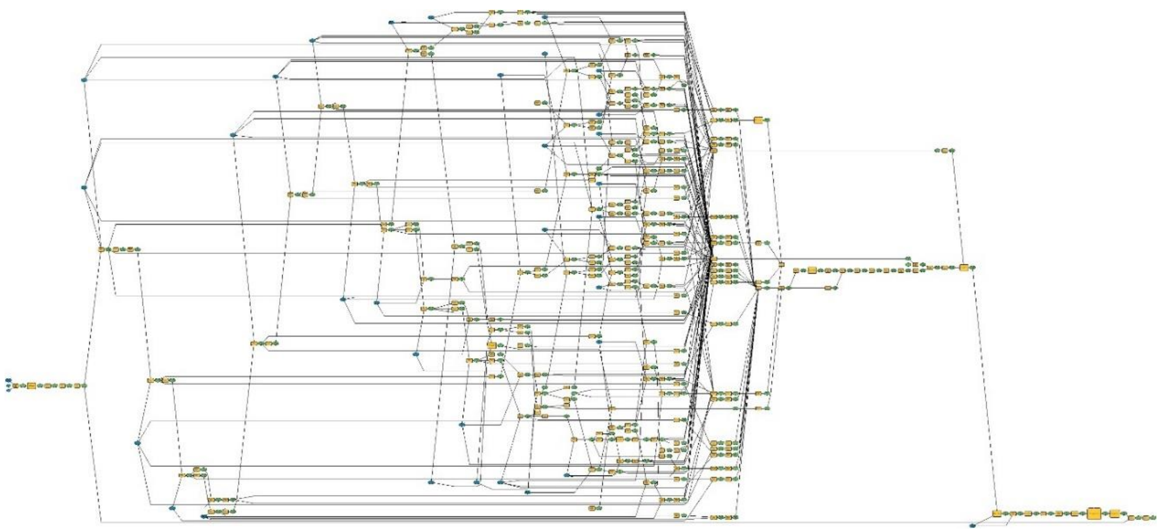


Figure S2.3 D) GSAAM Version 4: Multiple-decision tree model that incorporated all aspects of Version 3 and added functionality to specifically identify plume occurrences on Oil and Gas Fields and determine urban vs non-urban distinction.

Table S2.1 GSAAM V4 Tabular Product. This table displays the outputs that is produced by GSAAM V4. It includes 6 Categories: IPCC Level 1, IPCC Level 3, Vista-CA Facility, Vista-CA Sub-Facility, Oil and Gas Field, and Urban Oil and Gas Field. Vista-CA Facility, Sub-Facility, and Oil and Gas Field columns only produce associated Vista-ID's which can be linked back to individual infrastructure in the Vista-CA dataset.

IPCC_Level_1	IPCC_Level_3	Vista_CA_Facility	Vista_CA_Sub_Facility	Oil_Gas_Field	Urban_OG_Field
1. ENERGY	1A1 Energy Industries	REF000005	PWP000132		
1. ENERGY	1B2 Oil and Natural Gas	FLD000157		FLD000157	Yes
1. ENERGY	1B2 Oil and Natural Gas	FLD000157		FLD000157	Yes
1. ENERGY	1A1 Energy Industries	REF000013	PWP000229		
1. ENERGY	1A1 Energy Industries	REF000013	PWP000229		
4. WASTE	4A1 Managed Waste Disposal Sites	LNF000677			
4. WASTE	4A1 Managed Waste Disposal Sites	LNF000615			
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 & 3A2 Enteric Fermentation and Manure Management	DAF000259		FLD000326	No
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 & 3A2 Enteric Fermentation and Manure Management	DAF001728			
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 & 3A2 Enteric Fermentation and Manure Management	DAF001290		FLD000326	No
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 & 3A2 Enteric Fermentation and Manure Management	DAF000910			
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 & 3A2 Enteric Fermentation and Manure Management	DAF000416			
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 & 3A2 Enteric Fermentation and Manure Management	DAF001643			

Table S2.2 Vista-CA Overlap Analysis. This table shows the number of features that overlap with a given feature, which offers insight into the facility – sub-facility relations among different Vista-CA sectors, developing a hierarchy for GSAAM, and special spatial considerations (e.g. Oil and Gas Field Boundaries and their expansive coverage).

	Overlapping/Intersecting Vista-CA Sectors																
Number of ___ that Overlap/Intersect with >	Power Plants	Refineries	Compressor Stations	NG Fueling Stations	Oil and Gas Facilities	Oil and Gas Field Boundaries	Oil and Gas Wells	Processing Plants	Storage Fields	Feed Lots (100m)	Digesters	Dairies (100m)	Landfills	Composting Sites (100m)	Wastewater Treatment Plants	**TOTAL**	**Total Vista Elements**
Power Plants		16	81	0	16	85	43	3	1	0	1	2	45	7	26	326	433
Refineries	14		8	0	4	16	15	0	1	0	0	0	0	0	1	59	26
Compressor Stations	81	11		0	30	291	57	10	48	0	0	1	10	0	8	547	1120
NG Fueling Stations	0	0	0		0	32	0	0	1	0	0	0	5	3	0	41	208
Oil and Gas Facilities	18	28	30	0		3321	1060	27	40	0	0	0	42	0	4	4570	3356
Oi and Gas Field Boundaries	39	8	106	19	226		506	23	24	3	1	27	53	21	16	1072	516
Oil and Gas Wells	333	1291	111	0	5850	211735		1534	1185	0	0	2	2259	16	93	224409	225766
Processing Plants	4	0	9	0	18	25	23		12	0	0	0	0	0	0	91	26
Storage Fields	1	1	10	1	8	12	12	0		0	0	1	2	2	1	51	12
Feed lots (100m)	0	0	0	0	0	3	0	0	0		0	1	0	0	0	4	72
Digesters	1	0	0	0	0	2	0	0	0	0		31	0	0	0	34	33
Dairies (1000m)	2	0	1	0	0	64	2	0	3	1	31		0	2	0	106	1715
Landfills	41	0	7	5	18	87	80	0	2	0	0	0		54	0	294	714
Composting Sites (100m)	7	0	0	4	0	24	9	0	2	0	0	2	62		9	119	430
Wastewater Treatment Plant	22	1	7	0	3	14	15	0	1	0	0	0	0	9		72	149

VI. References

1. California Air Resources Board. Methane CH_4 Research Program. 2018. https://ww2.arb.ca.gov/our-work/programs/methane-research (accessed 6th September 2022).

2. California Department of Conservation. Online Data. 2019. https://www.conservation.ca.gov/dog/Online_ Data/Pages/Index.aspx. (accessed 23rd May 2019).

3. California Energy Commission. Supply and Demand of Natural Gas in California. 2019. https://www. energy.ca.gov/almanac/naturalgas_data/overview.html. (accessed 23rd May 2019).

4. Carranza, V., Rafiq, T., Frausto-Vicencio, I., Hopkins, F. M., Verhulst, K. R., Rao, P., Duren, R. M., and Miller, C. E. 2018. Vista-LA: Mapping methane-emitting infrastructure in the Los Angeles megacity, Earth Syst. Sci. Data, 10, 653-676, DOI: https://doi.org/10.5194/essd-10-653-2018.

5. Cusworth, D. H., Duren, R. M., Thorpe, A. K., Tseng, E., Thompson, D., Guha, A., Newman, S., Foster, K. T., Miller, C.E. 2020. Using remote sensing to detect, validate, and quantify methane emissions from California solid waste operations Environmental Research Letters. in press, DOI: https://doi.org/10.1088/1748-9326/ab7b99.

6. Cusworth, D. H., Duren, R. M., Yadav, V., Thorpe, A. K., Verhulst, K., Sander, S., Hopkins, F., Rafiq, T., Miller, C. E. 2020. Synthesis of methane observations across scales: Strategies for deploying a multitiered observing network. Geophysical Research Letters, 47, e2020GL087869, DOI: https://doi.org/10.1029/2020GL087869.

7. Cusworth, D. H., Jacob, D. J., Varon, D. J., Miller, C. C., Liu, X., Chance, K., Thorpe, A. K., Duren, R. M., Miller, C. E., Thompson, D. R., Frankenberg, C., Guanter, L., and Randles, C. A. 2019. Potential of next-generation imaging spectrometers to detect and quantify methane point sources from space. Atmosphere Measurement Techniques. 12 5655–5668, DOI: https://doi.org/10.5194/amt-12-5655-2019.

8. Duren, R. M., Thorpe, A. K., Foster, K.T., Rafiq, T., Hopkins, F. M. Yadav, V., Bue, B. D., Thompson, D. R., Conley, S., Colombi, N.K., Frankenberg, C., McCubbin, I. B., Eastwood, M. L., Falk, M., Herner, J. D., Croes, B. E., Green, R. O., Miller, C. E. 2019. California's Methane 1 Super-Emitters. Nature 575, 180–184. https://doi.org/10.1038/s41586-019-1720-3.

9. Duren, R. M., Thorpe, A. K., Sander, S. 2017. California Baseline Methane Survey Interim Phase 1 Report. California Air Resources Board.https://ww3.arb.ca.gov/research/methane/ca_ch4_survey_phase1_report_2017.pdf?_ga=2.87231512.1429954428.1558559590-776421153.1554761973. (accessed 23rd May 2019).

10. Eggleston, H. S., L. Buendia, K. Miwa, T. Ngara, and K. Tanabe. 2006. IPCC Guidelines for National Greenhouse Gas Inventories. Prepared by the National Greenhouse Gas Inventories Programme, vol. 4. IGES, Japan.

11. Elder, C. D., Thompson, D. R., Thorpe, A. K., Hanke, P., Walter Anthony, K. M., & Miller, C. E. 2020. Airborne mapping reveals emergent power law of Arctic methane emissions. Geophysical Research Letters, 47, e2019GL085707, DOI: https://doi.org/10.1029/2019GL08570 7.

12. Energy Information Administration Independent Statistics & Analysis. Layer Information for Interactive State Maps. 2019. https://www.eia.gov/maps/layer_info-m.php. (accessed 23rd May 2019).

13. Environmental Protection Agency. Facility Level Information on GreenHouse Gases Tool. 2017. https:// ghgdata.epa.gov/ghgp/main.do. (accessed 23rd May 2019).

14. Environmental Protection Agency. Geospatial Download Service. 2019. https://www.epa.gov/frs/geospatial -data-download-service. (accessed 23rd May 2019).

15. Environmental Protection Agency. Using GHG Inventory and GHGRP Data. 2019. https://cfpub.epa.gov/ghgdata/inventoryexplorer/data_explorer_flight.html (accessed 4th April 2020).

16. Etminan, M., G. Myhre, E. J. Highwood, and K. P. Shine. 2016. Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, Geophys. Res. Lett., 43, 12,614–12,623, DOI: doi:10.1002/2016GL071930.

17. Frankenberg, C., Thorpe, A. K., Thompson, D. R., Hulley, G., Kort, E. A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C., Conley, S., Bue, B. D., Aubrey, A. D., Hook, S., Green, R. O. 2016. Airborne methane mapping in the Four Corners area. Proceedings of the National Academy of Sciences. 113 (35) 9734-9739, DOI: 10.1073/pnas.1605617113.

18. Global Warming Potentials. Global Warming Potentials. United Nations, 2014. http://unfccc.int/ghg_data /items/3825.php (accessed 23rd May 2019).

19. Handan-Nader, C., Ho, D.E. 2019. Deep learning to map concentrated animal feeding operations. Nat Sustain 2, 298–306, DOI: https://doi.org/10.1038/s41893-019-0246-x.

20. Hopkins, F. M., Kort, E. A., Bush, S. E., Ehleringer, J. R., Lai, C.-T., Blake, D. R., and Randerson, J. T. 2016. Spatial patterns and source attribution of urban methane in the Los Angeles Basin, J. Geophys. Res. Atmos., 121, 2490–2507, DOI: doi:10.1002/2015JD024429.

21. Hopkins, F.M., Rafiq, T., and Duren, R.M. 2019. Sources of Methane Emissions (Vista-CA), State of California, USA. ORNL DAAC, Oak Ridge, Tennessee, USA, DOI: https://doi.org/10.3334/ORNLDAAC/1726.

22. Hulley, G. C., Duren, R. M., Hopkins, F. M., Hook, S. J., Vance, N., Guillevic, P., Johnson, W. R., Eng, B. T., Mihaly, J. M., Jovanovic, V. M., Chazanoff, S. L., Staniszewski, Z. K., Kuai, L., Worden, J., Frankenberg, C., Rivera, G., Aubrey, A. D., Miller, C. E., Malakar, N. K., Sánchez Tomás, J. M., and Holmes, K. T. 2016. High spatial resolution imaging of methane and other trace gases with the airborne Hyperspectral Thermal Emission Spectrometer (HyTES), Atmos. Meas. Tech., 9, 2393–2408, DOI: https://doi.org/10.5194/amt-9-2393-2016.

23. Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., and Petrescu, A. M. R. Submitted, 2017. Global Atlas of the three major Greenhouse Gas Emissions for the period 1970-2012, Earth System Science Data, essd-2017-79.

24. Jeong, S., Newman, S., Zhang J., Andrews. A. E., Bianco, L., Bagley, J., Cui, X., Graven H., Kim J., Salameh, P., LaFranchi, B.W., Priest, C., Campos-Pineda, M., Novakovskaia, E., Sloop, C. D., Michelsen, H. A., Bambha, R. P., Weiss. R. F., Keeling, R., and Fischer, M. L. 2016. Estimating methane emissions in California's urban and rural regions using multitower observations, Journal of Geophysical Research: Atmosphere. 121, 031–13,049, DOI: doi:10.1002/2016JD025404.

25. Jeong, S., Zhao, C., Andrews, A. E., Bianco, L., Wilczak, J. M., and Fischer, M. L. 2012. Seasonal variation of CH_4 emissions from central California, Journal of Geophysical Research: Atmosphere. 117, D11306, DOI: doi:10.1029/2011 JD016896.

26. Maasakkers J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., Hight, C., DeFigueiredo, M., Desai, M., Schmeltz, R., Hockstad, L., Bloom, A. A., Bowman, K. W., Jeong S., and Fischer, M. L. 2016.Gridded National Inventory of U.S. Methane Emissions. Environmental Science & Technology. 50, 13123-13133, DOI: doi: 10.1021/acs.est.6b02878.

27. Mehrotra, S., Faloona, I., Suard, M., Conley, S., and Fischer, M. L. 2017. Environ. Sci. Technol. 51, 12981−12987.

28. Mielke-Maday, I, et al. 2019. Methane source attribution in a U.S. dry gas basin using spatial patterns of ground and airborne ethane and methane measurements. Elem Sci Anth, 7: 13, DOI: https://doi.org/10.1525/elementa.351.

29. Miller, S.M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J. Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., Sweeney, C. 2013.Anthropogenic emissions of methane in the United States, Proc Natl Acad Sci USA. 110, 20018–20022, DOI: doi:10.1073/pnas.1314392110.

30. Nunez and Pavley. AB-32 Air pollution: greenhouse gases: California Global Warming Solutions Act of 2006. California Legislative Information. 2006. https://leginfo.legislature .ca.gov/faces/billNavClient.xhtml?bill_id=200520060AB32. (accessed 23rd May 2019).

31. Sheng, J.-X., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Zvala-Araiza, D., Hamburg, S. P. 2017. A high-resolution (0.1° × 0.1°) inventory of methane emissions from Canadian and Mexican oil and gas systems. Atmospheric Environment, Pergamon. www.sciencedirect.com/science/article/pii/S135223 1017301164?via%3Dihub.

32. Thompson, D. R., Leifer, I., Bovensmann, H., Eastwood, M., Fladeland, M., Frankenberg, C., Gerilowski, K., Green, R. O., Kratwurst, S., Krings, T., Luna, B., and Thorpe, A. K. 2015. Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane, Atmos. Meas. Tech., 8, 4383–4397, DOI: https://doi.org/10.5194/amt-8-4383-2015.

33. Thompson, D. R., Thorpe, A. K., Frankenberg, C., Green, R. O., Duren, R. M., Guanter, L., Hollstein, A., Middleton, E., Ong, L., and Ungar. S. 2016. Space-based remote imaging spectroscopy of the Aliso Canyon CH_4 superemitter, Geophys. Res. Lett., 43, 6571–6578, DOI: 10.1002/2016GL069079.

34. Thorpe, A.K., Duren, R., Conley, S., Prasad, K., Bue, B., Yadav, V., Foster, K., Rafiq, T., Hopkins, F., Smith, M., Fischer, M. L., Thompson, D., Frankenberg, C., McCubbin, I., Eastwood, M., Green, R. and Miller, C. E., 2020. Methane emissions from underground gas storage in California. Environ. Res. Letters, DOI: https://doi.org/10.1088/1748-9326/ab751d.

35. Thorpe, A., Frankenberg, C., Aubrey, A.D., Roberts, D., Nottrott, A.A., Rahn, T., Sauer, J.A., Dubey, M., Costigan, K.R., Arata, C., Steffke, A.M., Hills, S., Haselwimmer, C., Charlesworth, D., Funk, C.C., Green, R.O., Lundeen, S.R., Boardman, J.W., Eastwood, M., McFadden, J.P. 2016. Mapping methane concentrations from a controlled release experiment using the next generation airborne visible/infrared imaging spectrometer (AVIRIS-NG). Remote Sensing of Environment. 179. 104-115. 10.1016/j.rse.2016.03.032.

36. Thorpe, A.K., Frankenberg, C., Thompson, D.R., Duren, R.M., Aubrey, A.D., Bue, B.D., Green, R.O., Gerilowski, K., Krings, T., Borchardt, J. and Kort, E.A. 2017. Airborne DOAS retrievals of methane, carbon dioxide, and water vapor concentrations at high spatial resolution: application to AVIRIS-NG. Atmospheric Measurement Techniques, 10(10), 3833-3850, DOI: http://dx.doi.org/10.5194/amt-10-3833-2017.

37. Thurmond. AB-1496 Methane emissions (2015-2016). California Legislative Information. 2015. https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id=201520160AB1496.

38. Townsend-Small, A., S. C. Tyler, D. E. Pataki, X. Xu, and L. E. Christensen. 2012. Isotopic measurements of atmospheric methane in Los Angeles, California, USA: Influence of “fugitive” fossil fuel emissions, J. Geophys. Res., 117, D07308, Doi:10.1029/2011JD016826.

39. Yadav, V., Duren, R., Mueller, K.,Verhulst, K. R., Nehrkorn, T., Kim, J., Weiss, R.F., Keeling, R., Sander, S., Fischer, M. L., Newman, S., Falk, Matthias, Kuwayma, T., Hopkins, F., Rafiq, T., Whetstone, J., and Miller, C. 2019. Spatio-temporally resolved methane fluxes from the Los Angeles megacity. Journal of Geophysical Research: Atmosphere.124, DOI: https://doi.org/10.1029/2018JD030062.

40. Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L. 2012. On the Sources of Methane to the Los Angeles Atmosphere. *Environmental Science & Technology., 46* (17), 9282-9289, DOI: https://doi.org/10.1021/es301138y.

41. Zavala-Araiza, D., Lyon, D. R, Alvarez, R., Davis, K., Harriss, R., Herndon, S., Karion, A., Kort, E. A., Lamb, B., Lan, X., Marchese, A. J., Pacala, S. W., Robinson, A. L., Shepson, P., Sweeney, C., Talbot, R., Townsend-Small, A., Yacovitch, T., Zimmerle, D., Hamburg, S. 2015. Reconciling divergent estimates of oil and gas methane emissions. Proceedings of the National Academy of Sciences of the United States of America. 112. 10.1073/pnas.1522126112.

42. Zhao, C., Andrews, A. E., Bianco, L., Eluszkiewicz, J., Hirsch, A., MacDonald, C., Nehrkorn T., and Fischer, M. L. 2009. Atmospheric inverse estimates of methane emissions from Central California, Journal of Geophysical Research,114, D16302, DOI: doi:10.1029/2008JD011671.

Chapter 3

Top-Down and Bottom-Up Methane Emissions Development and Assessment for Kern County, California

I. Introduction

California's Kern County contains a plethora of methane (CH_4) emitting infrastructure which constitute a significant portion of the state's CH_4 emissions budget. Kern County contains the largest number of oil and gas infrastructure in the state of California (~90% of the state's total).[4] Proximity to CH_4 oil and gas sources poses a health risk to the nearly 1 million people that reside and work in Kern County. Consequently, California has committed to reductions in CH_4 emissions and enacted legislation (SB 1383) to achieve its climate goals of significantly reducing greenhouse (GHG) emissions (Assembly Bill 32) by 2030.[1, 2, 3] However, uncertainty in the quantification and source apportionment of California CH_4 emissions hinders effective policy planning and mitigation implementation along with the substantiation of emission reductions for the future. Additionally, a lack of fine-scale CH_4 emission assessments for this region necessitates further investigation and development of robust CH_4 emission products.

There have been two main methods of quantifying CH_4 emissions from a large geographic area such as Kern County: top-down observations and bottom-up modeling. Top-down and bottom-up emission estimates allow for the estimation of CH_4 in a given area and can be used complementary to one another. Bottom-up approaches require utilizing specific activity data and local processes to develop models that can estimate and scale emissions across sectors. Top-down approaches measure patterns through direct atmospheric observation of various CH_4 sources and sinks. Both top-down and bottom-up methods contain their own advantages and deficiencies. Discrepancies in these methods make it more difficult to quantify emissions and spatially attribute them accurately.

Downscaled versions of state-wide inventories have been found to consistently underestimate emissions compared to top-down atmospheric studies.[10, 11, 12, 13, 16] Top-down emission evaluations show that CH_4 emissions for California are underestimated by ~50% or more depending upon the area.[8, 9, 10, 11] With regards to spatial context, existing top-down studies aggregate CH_4 emission estimates, while bottom-up inventories are often too spatially coarse to identify individual sources from facilities.[16, 17, 18] These differences between top-down and bottom-up emissions estimates indicate that the understanding of CH_4 sources and overall emissions can be incomplete in some areas. This incompleteness in turn can potentially yield uncertainty in the application of appropriate policy measures that are designed to mitigate CH_4 emissions accordingly.

We explored a complementary method of contextualization by developing and utilizing both top-down and a bottom-up approaches. We hypothesized that there is a greater impact on CH_4 emissions from the oil and gas sector that has been previously unaccounted or underrepresented in publicly available CH_4 emissions datasets for Kern County. This research builds a foundation for future studies measuring and modeling CH_4 emissions produced from the considerable number of oil and gas infrastructure and facilities in this region. This study shows how the impacts from under accounting in the oil and gas sector resulted in significant underestimation of CH_4 emissions in the CARB, EPA, and CALGEM products. Here we present two approaches for constraining and estimating CH_4 emissions in California's Kern County: 1) A top-down facility-level measurement attributed to CH_4 emitting infrastructure and 2) A high-resolution bottom-up facility-level model developed from validated geospatial data. Tracking emissions at the facility level is valuable since mitigation is most actionable at this scale. We expand on previous work and use the end-products generated from Rafiq et al.'s Vista-CA Geospatial Source Attribution Automated Model (GSAAM) (2020) that attributes Duren et al.'s Airborne Visual Imaging Infrared Spectrometer – Next Generation (AVIRIS-NG) (2019) emissions survey data (https://data.carbonmapper.org) to Vista-California (Vista-CA) geospatial data (https://doi.org/10.3334/ORNLDAAC/1726) in order to systematically estimate and scale CH_4 emissions in Kern County.[4, 19] Using this dataset, we analyzed CH_4 emission trends at the facility and Intergovernmental Panel on Climate Change (IPCC) sectoral levels and compared them with the California Air Resources Board's Pollution Mapping Tool (CARB PMT) facility-level data (https://ww3.arb.ca.gov/ei/tools/pollution_map/)[15], the California Greenhouse Gas Emissions Measurement (CALGEM) Project bottom-up inventory (http://calgem.lbl.gov/about.html)[16], the Environmental Protection Agency's EPA GHG bottom-up Inventory (Maasakkers' EPA) (https://www.epa.gov/ghgemissions/gridded-2012-methane-emissions#data)[17], and the publicly available EPA geospatial GHG emissions inventory: Facility Level Information on GreenHouse gases Tool (EPA FLIGHT) (https://ghgdata.epa.gov/ghgp/main.do#)[20]. Finally, we contextualize how our estimates compare with those from the aforementioned studies in order to assess the consistency and relevancy of our approach.

II. Methods

We collected, processed, and generated 6 different CH_4 emissions datasets for Kern County. First, we utilized geospatial CH_4 inventories, specifically the Vista-California database (Vista-CA). Vista-CA is a geospatial database of 901,009 validated elements of potential CH_4 emitting infrastructure developed from publicly available datasets that have been validated and standardized for the entire state of California.[4, 22] This database includes 17 CH_4 source layers that were categorized into facilities and sub-facilities and is organized according to the CARB GHG Inventory, which itself is based on the framework established by the Intergovernmental Panel on Climate Change (IPCC).[23, 24] This systematic approach for understanding the distribution of CH_4 sources was demonstrated by Carranza et al. for the Los Angeles (LA) Basin and Rafiq et al. for California through the development of Vista-CA.[4, 21]

We further employed two policy tools used for tracking GHG emissions for this study. The first was the U.S. Environmental Protection Agency Facility Level Information on GreenHouse gases Tool (EPA FLIGHT: https://ghgdata.epa.gov/ghgp/main.do).[20] EPA FLIGHT contains a geospatial database of approximately 8,000 facilities that report annually to the EPA Greenhouse Gas Reporting Program (GHGRP). EPA FLIGHT tracks facilities that emit more than 25,000 $MTCO_2$ eq a^{-1}, and accounts for 85-90% of emissions included in the official EPA GHG Reporting Program. We collected EPA FLIGHT CH_4 data for 2016, 2017, and 2018 which contained a geospatial dataset of around 400 CH_4 reporting facilities for California. EPA FLIGHT differs from the official EPA GHG inventory because they are based on reported emissions at the facility scale, not activity data. EPA FLIGHT data was manually binned into Vista-CA categories based either on direct spatial or tabular matches with Vista-CA facilities or publicly available facility records. The second tool was the emissions data from California Air Resources Board's Pollution Mapping Tool (CARB PMT) (https://ww3.arb.ca.gov/ei/tools/pollution_map/).

CARB PMT tracks GHG and criteria pollutant emissions from large facilities in California that are subject to the GHG Mandatory Reporting Regulation.[16]

Additionally, we utilized the 2016 California Greenhouse Gas Emissions Measurement (CALGEM) Project bottom-up gridded product for California for this project.[6, 8, 16] CALGEM computes emissions for landfills, dairy and non-dairy livestock, wastewater treatment plants, natural gas and petroleum production/local processing, and petroleum refining and mobile sources. It provides gridded emissions at annual average surface fluxes in units of 1×10^{-9} mol m^{-2} s^{-1} at $0.1° \times 0.1°$ spatial resolution. CALGEM's livestock and non-livestock data incorporated the 2012 county-level dairy statistics from the U.S. Department of Agriculture. CALGEM utilized natural gas well data from California's Department of Conservation (CDC) and estimated gas well emissions using production data while assuming a leakage rate of 1%.[6, 8, 16] Additionally, they utilized data from the 2012 California Air Resources Board (CARB) state totals for the rest of their anthropogenic and fossil sectors including petroleum refining and wastewater treatment plants. CARB borrows their bottom-up methods for calculating sectoral CH_4 emissions from the 2006 IPCC report (Supplement S4.1).

We also obtained the 2016 EPA GHG Inventory developed by Maasakkers et al.[17] Maasakkers et al. developed a gridded anthropogenic CH_4 emissions inventory of the U.S. at a $0.1° \times 0.1°$ spatial resolution using units of molecules $cm^{-2}s^{-1}$.[17] They utilized 2012 CH_4 emissions from the 2014 EPA Greenhouse Gas Inventory (GHGI) which includes detailed calculation methods to estimate emissions from all CH_4 relevant sectors

(Supplement S4.2). CALGEM and Maasakkers' EPA calculation methods are detailed in the supplement to this work (Supplement S3).

Advances in top-down measurements have allowed the detection of CH_4 point sources from surveys of areas spanning 1,000 to 100,000 km^2 to be detected at sub-meter scales through imaging spectroscopy. We used these CH_4 plume observations from the California survey conducted using the Airborne Visible/Infrared Imaging Spectrometer - Next Generation (AVIRIS-NG) instrument.[19] AVIRIS-NG is capable of detecting concentrated CH_4 plumes by measuring ground-reflected solar radiation across 427 contiguous spectral bands ranging from 350 to 2,500 nm wavelengths with 5 nm spectral sampling at 1 - 3 m spatial resolution.[25, 26, 27] The CH_4 retrieval is based on absorption spectroscopy between 2,100 and 2,500 nm and provides a mixing ratio length that represents a CH_4 enhancement in parts per million-meter (ppm-m). AVIRIS-NG's spectral resolution, high spatial resolution, and high signal-to-noise ratio has permitted high-resolution mapping of CH_4 as well as CO_2 and H_2O. AVIRIS-NG has consistently detected and quantified CH_4 point sources from multiple emissions sectors for emissions as small as 2-10 CH_4 kg hr^{-1}, depending on surface albedo and aircraft/ground speed. Through this system, Duren et al. surveyed 60% of Vista-CA CH_4 emitting infrastructure in California and found that a few hundred CH_4 point sources accounted for 34-46% of the overall statewide CH_4 budget.[4, 19, 22] AVIRIS-NG is able to detect fugitive CH_4 point sources with high precision but is unable to quantify area sources or sources that are spread out over a large scale (km^2). Consequently, sectors with large footprints such as landfills are more difficult to integrate using AVIRIS-NG measurements.

In this study we used the 2,424 AVIRIS-NG CH_4 plume observations collected during the 2016 – 2018 California Methane Campaign for facility-level attribution across Kern County. All data presented comes from the surface coverage of the AVIRIS-NG campaign that surveyed approximately 16% of Kern County by area. We utilized Vista-CA Geospatial Source Attribution Automated Model's (GSAAM) attribution results of the AVIRIS-NG data against the Vista-CA dataset for facility-based emissions assessment in this study.[4] GSAAM enables automated attribution of Vista-CA facility and sub-facility to point locations of CH_4 plumes using a series of structured geospatial relationships. Further details on the methods of this process can be found in Rafiq et al. 2020. The facility attribution of individual AVIRIS-NG plumes with Vista-CA enabled emissions categorization and comparison with EPA FLIGHT facility emissions.

After attributing observed plumes to each facility, we reanalyzed the data to account for overflights of emitting facilities where no plumes were seen. While AVIRIS-NG is a high spatial resolution product, it only provides a snapshot of emissions during sporadic overflights. In order to account for temporal gaps in observation from AVIRIS-NG into our emissions calculation, we developed data that identifies all plumes flown and all flightlines that flew over those plumes, and for each plume, whether that flightline that flew over it, observed that given plume or did not observe that given plume (Supplement S3). This CH_4 emissions source data resulted in a total of 9,482 plume-to-flightpath spatial connections stemming from the 795 unique sources as identified by AVIRIS-NG. We ran the finalized source data product through Vista-CA GSAAM to obtain Vista-CA facility and sub-facility spatial attributions, along with explicit identification of Oil and Gas Field spatial attributions for all 9,482 plume-to-flightpath spatial connections. We have taken the total number of times a given facility in each study area was flown by AVIRIS-NG and recorded the number of times it was observed to be emitting and observed to not be emitting in order to utilize our methods for emissions estimation (Tables S3.1 and S3.2).

Across Kern County, over 2,000 facilities surveyed by AVIRIS-NG were observed to be emitting (Table S3.1).[19] Because of the CH_4 sensitivity limitations of the AVIRIS-NG instrument we developed an emissions constraint. This was based on the notion that each facility that was flown by AVIRIS-NG but not observed to be emitting any CH_4 could potentially be emitting CH_4 below the sensitivity limits of the AVIRIS-NG instrumentation of 2-10 kg CH_4 hr^{-1}.[19] For such a given facility where AVIRIS-NG did not observe any emissions, we applied two different measures for constraining potential emissions: a randomized estimate and the maximum bound. We took AVIRIS-NG's sensitivity limits into account by assigning estimates that ranged from 0-10 kg CH_4 hr^{-1}. For the random assignment, all non-observationally emitting facilities were assigned a randomly generated number between 0-10 kg CH_4 hr^{-1}. For the maximum bound, we assigned the highest sensitivity value of 10 kg CH_4 hr^{-1} to all non-observationally emitting facilities. Both observed and flown facilities and non-observed but flown facilities that belonged to a given Vista-CA category were then summed together for both the random assignment and the maximum bound. This method of sensitivity consideration did not result in any significant differences to facility, sectoral, or regional emissions estimates and thus was not further utilized. We used the direct observational AVIRIS-NG data to develop our results.

The direct observational data method utilizes the sum of all fugitive and hotspot plume emissions from a given facility and multiplies it by a persistence measurement which is the ratio of the observed plumes to the total number of overflights. Normalizing by the persistence allows for us to potentially account for various temporal characteristics in facility observations.

Equation 1.

$$\boldsymbol{Q_{TU}} = \sum(\boldsymbol{\bar{Q}_n} \times \boldsymbol{P_n}), \qquad \boldsymbol{where\ \ P} = \left(\frac{\boldsymbol{O}}{\boldsymbol{T}}\right)$$

Where $\boldsymbol{Q_{TU}}$ is the total facility emissions estimate, $\boldsymbol{Q}$ is the individual plume estimated by AVIRIS-NG,[19] $\boldsymbol{P}$ is the persistence, $\boldsymbol{n}$ is the source number, $\boldsymbol{O}$ is the total number of observed plumes, and $\boldsymbol{T}$ is the total number of overflights for a given facility.

In order to generate bottom-up emission estimates we employed a combination of methods ranging from obtaining facility numbers from CARB and EPA, generating our own numbers using IPCC bottom-up equations, or through adapting existing bottom-up models. We developed emissions for 14 Vista-CA categories which include composting sites, dairies, digesters, feed lots, landfills, natural gas stations, oil and gas facilities, oil and gas fields, oil and gas wells, pipelines, power plants, processing plants, refineries, and wastewater treatment plants. Using the EPA's GHG inventory for natural gas systems, we employed emission factors and the IPCC tier 1 methodology for the following sectors: natural gas stations, oil and gas facilities, oil and gas wells, pipelines, and processing plants. We adopted Maasakkers' EPA modeled CH_4 data for the following sectors: composting sites, landfills (portion of the data was obtained from CARB), and wastewater treatment plants. Furthermore, we leveraged Marklein et al.'s work for computing emissions for dairies, digesters, and feedlots, employed Vafi et al.'s modeled emissions for oil and gas fields, and utilized Rafiq et al.'s study for modeled power plant emissions (Vafi et al. 2021). Lastly, we utilized CARB's reported facility CH_4 emissions data for refineries. Vista-CA Bottom-Up data was gridded to 1 km by 1 km and 10 km by 10 km. Detailed information for Vista-CA Bottom Up CH_4 emissions generation can be found in the supplementary document (Supplement S2).

We categorized CH_4 emissions based on biogenic and fossil signal types among all of the datasets by binning facilities to those two categories. Biogenic type emissions relate to emissions from the anerobic process or decay of matter which includes the following facilities: composting sites, dairies, digesters, feed lot, landfills, and wastewater treatment plants. Fossil type emissions relate to emissions originating from the process of using, distributing, processing, and the creation of fossil fuels, and includes the following facilities and sub-facilities: distribution pipelines, natural gas fueling stations, natural gas stations, oil and gas facilities, oil and gas fields, power plants, processing plants, refineries, and storage fields. For comparative purposes, CARB PMT, EPA FLIGHT, CALGEM, Maasakkers' EPA, AVIRIS-NG Source Data, and Vista-CA Bottom-Up emissions were converted to Gg of CH_4 a^{-1} and all gridded emissions products were gridded to 10 km by 10 km (Figures 3.1 and 3.5). Lastly, emissions from all 6 datasets were apportioned and normalized to IPCC level 3 source categories: 1A1 Energy Industries, 1B2 Oil and Natural Gas, 3A1 and 3A2 Enteric Fermentation and Manure Management, 4A1 Managed Waste Disposal, 4B Biological Treatment of Solid Waste, and 4D1 & 4D2 Domestic and Industrial Water Treatment and Discharge (Figure 3.5).

III. Results

We present here two systematic methods and products: 1) AVIRIS-NG Source Data: top-down accounting of flight line coverages of categorical fugitive emissions towards facility estimates and 2) Vista-CA Bottom-Up modeling derived from a combination of activity data, emission factors and existing facility-level geospatial emissions datasets. Using the IPCC methodology along with ancillary datasets for bottom-up emission modeling, we generated a 10 km x 10 km grid that encompasses 14 Vista-CA layers for Kern County (Figure 3.1A). The majority of emissions are concentrated in the western half of the county which is also evidenced by other estimates (Figures 3.2 and 3.3). The majority of the hot spots (red grid cells) in both Vista-CA Bottom-Up and AVIRIS-NG Source Data come from dairy farms and oil and gas related infrastructure. There exists a heavy influence of CH_4 emissions from the western border where an enormous amount of oil and gas infrastructure exists, specifically from the Midway-Sunset oil and gas field (Figures 3.2 and 3.4). The cluster in the center of the county is located near the City of Bakersfield and the long line extended over into the southeast is indicative of the pipeline infrastructure (Figure 3.2). Highly concentrated CH_4 emissions are exemplified in Vista-CA Bottom-Up's 1 km pixel which is at least 5 times larger than the highest concentrated 10 km pixel of the other datasets (Figures 3.1 and 3.3).

Vista-CA contains more records for facilities than other publicly accessible geospatial products and has attributed and generated emissions for close to 150,000 infrastructure elements and facilities in Kern County (Figures 3.2 and 3.3). Vista-CA Bottom-Up's total CH_4 emissions estimate for Kern County was 133.18 Gg CH_4 a^{-1}, which was 30% higher than total estimates from CALGEM and lower than estimates from Maasakkers' EPA. Vista-CA Bottom-Up's four highest emitting sectors were dairies (59.03 Gg CH_4 a^{-1}), oil and gas fields (43.91 Gg CH_4 a^{-1}), natural gas stations (12.86 Gg CH_4 a^{-1}), and landfills (10.71 Gg CH_4 a^{-1}), which constituted 96% of the total emissions from Kern County (Table S3.1).

The AVIRIS-NG Source Data product contains 79 persistence normalized CH_4 facility estimates which were derived from 631 detected plumes from dairies, landfills, natural gas stations, oil and gas facilities, oil and gas fields, power plants, processing plants, refineries, and wastewater treatment plants (Figure 3.4). AVIRIS-NG surveyed over 90% of the entire Kern County Vista-CA infrastructure dataset, of which does not even account for the repeated flights conducted over infrastructure dense and CH_4 hotspots. AVIRIS-NG Source Data's total CH_4 emissions estimate for Kern County of 130.1 Gg CH_4 a^{-1} was comprised of a majority fossil signal, predominantly from oil and gas infrastructure (Figures 3.4 and 3.5B). This estimate was 30% higher than CALGEM and lower than Maasakkers' EPA. The AVIRIS-NG Source Data product accounts for around 98% of the entire Kern County CH_4 budget as compared with Vista-CA Bottom-Up.

It is evident that there is general spatial agreement of the distribution of CH_4 emissions among all datasets in this study (Figures 3.1A-3.1E and 3.4). The intensity of emissions is captured by Vista-CA Bottom-Up, Maasakkers' EPA, and CALGEM. CARB PMT and EPA FLIGHT are similar to one another, but their distributions seem to not capture the majority of the CH_4 signal emanating from these different oil and gas and dairy infrastructures. Furthermore, the magnitude of emissions is highest in the Maasakkers'

EPA. dataset. CALGEM, AVIRIS-NG Source Data, and Vista-CA Bottom-Up's overall magnitudes are very similar while CARB PMT and EPA FLIGHT account for less than 4% of the total CH_4 emissions from Kern County (Figure 3.5A). While CALGEM has the smallest maximum grid cell value (1.65 Gg CH_4 a^{-1}), the maximums are similar between CARB PMT (2.76 Gg CH_4 a^{-1}), EPA FLIGHT (2.26 Gg CH_4 a^{-1}), and Maasakkers' EPA (2.56 Gg CH_4 a^{-1}) (Figure 3.1B – 3.1E). Vista-CA Bottom-Up's maximum grid cell magnitude is more than 7 times higher than any of the other datasets presented when viewed within the 10 km x 10 km summation grid (Figure 3.1A). When looking at the magnitude of the emissions across the data, Maasakkers' EPA total emission was the highest by a wide margin at 274.5 Gg CH_4 a^{-1} (Table S3.1). CALGEM, AVIRIS-NG and Vista-CA Bottom-Up were more in line with each with an average of 121.3 Gg CH_4 a^{-1} (Figure 3.5A).

There is some evidence of agreement in the CH_4 emissions assessment of biogenic and fossil type emissions across these datasets (Figure 3.5B). Because CARB PMT does not contain any data on biogenic infrastructure, the entire signal was fossil. Maasakkers' EPA and AVIRIS-NG Source Data ratio of biogenic emissions to fossil emissions constituted a majority fossil signal at 70% and 92%, respectively (Figure 3.5B). The opposite was evident for both CALGEM and EPA FLIGHT as biogenic related emissions signaled the majority with 70% and 86%, respectively (Figure 3.5B). Vista-CA Bottom-Up had a near even split between both with a slight edge to biogenic emissions by about 3% (Figure 3.5B). EPA FLIGHT's entire biogenic signal was derived from 2 landfills. Additionally, the magnitude of biogenic emissions in Vista-CA Bottom-Up, Maasakkers'

EPA and CALGEM were very similar at 72.5 Gg CH_4 a^{-1}, 82.4 Gg CH_4 a^{-1}, and 69.2 Gg CH_4 a^{-1}, respectively (Table S3.1 and S3.2).

Source apportionment of CH_4 was determined by proportionally normalizing emissions into IPCC Level 3 categories to enable comparison among the datasets (Figure 3.6). EPA FLIGHT contained emissions for 3 out of the 6 IPCC Level 3 categories and is dominated by Managed Waste Disposal (4A1) category while CARB PMT only contained emissions for 2 out of the 6 categories which comprised of Energy Industries and Oil and Natural Gas (1A1 and 1B2) at 6% and 94%, respectively (Figure 3.6). Vista-CA Bottom-Up's Oil and Natural Gas (1B2) and Enteric Fermentation and Manure Management (3A1 & 3A2) categories were the highest and pretty even at around 45% with the next highest category being Managed Waste Disposal (4A1) at 8%. AVIRIS-NG Source Data and Maasakkers' EPA identified 1B2 Oil and Natural Gas (90% and 70% respectively) as their largest emitting category with 3A1 & 3A2 Enteric Fermentation & Manure Management as their second largest category (5.2% and 24%, respectively) (Figure 3.6). The opposite was modeled with CALGEM data with 64% of the emissions coming from the 3A1 & 3A2 category while around 10% came from the 1B2 category and CALGEM is the only dataset with significant signals coming from the 4B and 4D1 & 4D2 categories (Figure 3.6).

IV. Discussion

The Vista-CA Bottom-Up and AVIRIS-NG Source Data products have provided high resolution CH_4 emissions of Kern County that are comparable if not improved to existing inventories. The AVIRIS-NG Source Data product enables a more thorough source apportionment accounting of CH_4 based on actual measurements that is likely to be more robust than simple activity and emission factor methods that do not capture fugitive or anomalously large sources that are thought to be common for CH_4, especially in California. There was general agreement in biogenic and fossil emission characterization among AVIRIS-NG Source Data and Maasakkers' EPA suggesting that the methods used to generate and scale CH_4 emissions for AVIRIS-NG Source Data were consistent and robust for source apportionment analysis. Vista-CA was better paired with CALGEM as the biogenic signal seemed to be a little greater but was not as stark as with AVIRIS-NG or Maasakkers' EPA While biogenic and fossil proportions across all datasets illustrate the differences in source apportionment of CH_4 emissions, they also highlight that the majority of the signal is constituted from two sectors: dairy and oil and gas. With this in mind, this further demonstrates the notion that a small number of super-emitters that encompass a small surface area are responsible for the majority of Kern County's CH_4 budget.[19]

Dairy emissions were fairly consistent among CALGEM, Maasakkers' EPA and Vista-CA Bottom-Up showing contributions close to 60 Gg CH_4 a^{-1}. This suggests that the dairy sector is producing on average at least 45% of the emissions for Kern County from only 44 facilities (Figure 3.5B). This is significant as it not only exemplifies Duren et al.'s super-emitter characteristic but potentially enables a more targeted approach to curb and mitigate CH_4 emissions.[17] Measured emissions from AVIRIS-NG's survey of 35 dairies showed about a 5% contribution to the total versus a 45% contribution to the total budget from the Vista-CA Bottom-Up modeled data. This indicates the need for improved techniques for observing dairy emissions in the field as AVIRIS-NG is possibly underestimating emissions from this sector due to a few factors. Some of these factors include inability to persistently monitor a given facility, changes in operations over a given period of time, and seasonal changes across the dairy facility. Dairies were not included in CARB PMT and EPA FLIGHT which highlights critical need to account for them in their reported emission products (Table S3.2).

Over 90% of the emissions in the observation-based AVIRIS-NG Source Data came from the oil and gas sector. In the Vista-CA Bottom-Up dataset, the oil and gas sector contributed around 45% towards the total Kern County CH_4 emissions budget. Maasakkers' EPA oil and gas contribution of 70% is in the middle of AVIRIS-NG Source Data and Vista-CA Bottom-Up. CALGEM, CARB PMT, and EPA FLIGHT show significant differences in the oil and gas emissions budgets. These results suggest that the oil and gas infrastructure in Kern County is not completely quantified nor is it completely accounted for. Furthermore, even with proper accounting of CH_4 emission sources from this sector, emissions have not been quantified accordingly as is evident by the AVIRIS-NG top-down observations of oil and gas wells, oil and gas fields, and oil and gas facilities. It is also evident that the number of facilities that are being tracked in these datasets are incomplete as evidenced from the Vista-CA Bottom-Up dataset. This is critical due to the high number of fugitive CH_4 emissions observed from this sector. We find the oil and gas sector is underestimated in public inventories and datasets. Even the modeled and observed measurements we developed (Vista-CA Bottom-Up and AVIRIS-NG Source Data) could potentially be under accounting emissions since Vista-CA does not track other possible fugitive sources such as well pump jacks, gathering lines and ancillary structures and

AVIRIS-NG could have missed emissions that were below its sensitivity threshold of 2 kg CH_4 h^{-1}.

High concentrations of CH_4 emissions from oil and gas infrastructure were also observed in another independent top-down study by NASA's CO_2 and Methane Experiment (COMEX) team.[30] COMEX looked at infrastructure and facilities located in the Kern River oil and gas field, Kern Front oil and gas field, and Poso Creek oil and gas field.[30] For example, in 2014 they detected leaks in oil and gas wells in the Poso Creek Oil field emitting point source concentrations greater than 1,000 kg CH_4 hr^{-1} and was corroborated by complementary AVIRIS-NG data at the time.[30] Similarly, between 2016 – 2018, AVIRIS-NG detected 67 CH_4 plumes within Poso Creek Oil field with emissions ranging from 41 - 724 kg CH_4 hr^{-1}. Our AVIRIS-NG Source Data product estimates persistence normalized emissions from all AVIRIS-NG detections within Poso Creek Oil field from 2016-2018 to be 187.6 kg CH_4 hr^{-1}. Both the COMEX study and AVIRIS-NG Source data suggest that a small number of fugitive leaks from faulty infrastructure can disproportionately contribute to CH_4 emissions from oil and gas related operations.[19] Fugitive leaks, which are unpredictable, can be difficult to account for in bottom-up modeling proposing that actual emissions estimates could be higher than even modeled estimates.

The lack of accounting in the oil and natural gas sector is greatly exemplified in Kern County with around 90% of the total potential CH_4 emitting facilities and infrastructure coming from this sector alone and neither CARB PMT nor EPA FLIGHT contain at least 1% of the records from the IPCC Level 3 1B2 category as compared to the

Vista-CA CH_4 dataset.[4, 20] This explains the discrepancy in source apportionment and magnitude of emissions for CARB PMT and EPA FLIGHT Kern County emissions. As described earlier, EPA FLIGHT tracks facilities that emit more than 25,000 $MTCO_2$ eq a^{-1} which means that any facilities that do not meet this threshold are not included in their database (25 facilities in Kern County) and CARB PMT tracks pollutants from large facilities that are subjected to their Mandatory Reporting Regulations (43 facilities Kern County), which also means that smaller facilities that are not subjected to this program are not included (Table S3.2). This suggests both of their bottom-up data collection methodologies are potentially missing facilities and sectors altogether that do significantly contribute to the overall CH_4 budget for Kern County (2,018 facilities for the region) (Table S3.1). As a result, reported emissions are not a good proxy for observed fugitive CH_4 emissions nor are they are they near modeled emission estimates. Vista-CA contains records for close to 150,000 CH_4 emitting facilities and infrastructure in Kern County, which is several magnitudes higher than any available geospatial dataset for this region (Tables S3.1 and S3.2).

The method we present here for developing the AVIRIS-NG Source Data product focuses primarily on point sources, fugitive emissions, and hot spots and thus it is able to reveal super-emitter activity. Yet, persistent observations are difficult to complete using the top-down AVIRIS-NG instrument as it flies and completes surveys for a certain period of time. These gaps in temporal coverage make it difficult to assess the persistence of high or anomalous CH_4 sources. For example, dairy facilities operate differently throughout the day and year, and to track these changes requires a more persistent technique closer to the actual facilities which the AVIRIS-NG instrument is not completely ideal for. Moreover, there are inherent limitations in employing CH_4 emissions calculations based on activity data and emission factors as was done for the Vista-CA Bottom-Up product. The accuracy of this technique is dependent on the accuracy and specificity of the underlying emission factors. Specific emission factors must be validated over time and requires thorough technical evaluation. Validity of emission factors and activity data relies on the age of the data and where it is sourced from as estimates of CH_4 emissions could be unintentionally calculated as higher or lower than what is actually observed based on these two factors. Older emission factors may not include updates in technology or adopt newer procedures that could potentially result in an incorrect emission characterization. Utilizing the most up to date emission factor and activity data can be challenging since verification of data and accessibility to data vary.

There exists a clear gap in emission estimates among the data that we have developed for Kern County and the data we have obtained from these publicly available ancillary sources. The existing geospatial inventories and bottom-up models underestimate the signals from the dairy and oil and gas sectors, they are unable to capture the influence of super-emitters, and are insufficient in providing high-resolution facility level scales for characterization of sources. The explanatory capability of the Vista-CA Bottom-Up and AVIRIS-NG Source Data emissions products provide complementary and comparable source apportionment information along with geospatial characteristics. Accounting for nearly all potential CH_4 emitters, allows for better constraining of emissions and improves the geospatial understanding of sources. Using a combination of bottom-up constraints and top-down attribution can produce valuable results and there is a clear need for additional bottom-up and top-down CH_4 emissions research for Kern County

Overall, these findings suggest that Kern County CH_4 emission trends observed in the AVIRIS-NG Source Data and modeled in Vista-CA Bottom-Up are in line with or exceed other publicly available bottom-up emission estimates. The trends at the facility and IPCC category scales require further analysis and this work can be further utilized to develop a more spatially resolved bottom-up emissions product to inform an emissions prior for inverse modeling analysis. This study provides a foundation for potentially extending these bottom-up and top-down methodologies to other critical areas across the state and country that are affected by CH_4 and towards augmenting future work addressing climate change. The City of Bakersfield in Kern County has a large population of close to 400,000 people and is adjacent to half a dozen oil and gas fields, dozens of dairy farms, hundreds of oil and gas facilities, and thousands of oil and gas wells, all within a few kilometers. Higher CH_4 emissions in proximity to these populated areas in Kern County opens up the potential for harmful health consequences and further research will be needed to assess these dynamics. Ultimately, addressing emissions at finer scales can yield meaningful actionable policies and productive monitoring procedures that are beneficial for these communities that live in proximity to these CH_4 emitting structures.

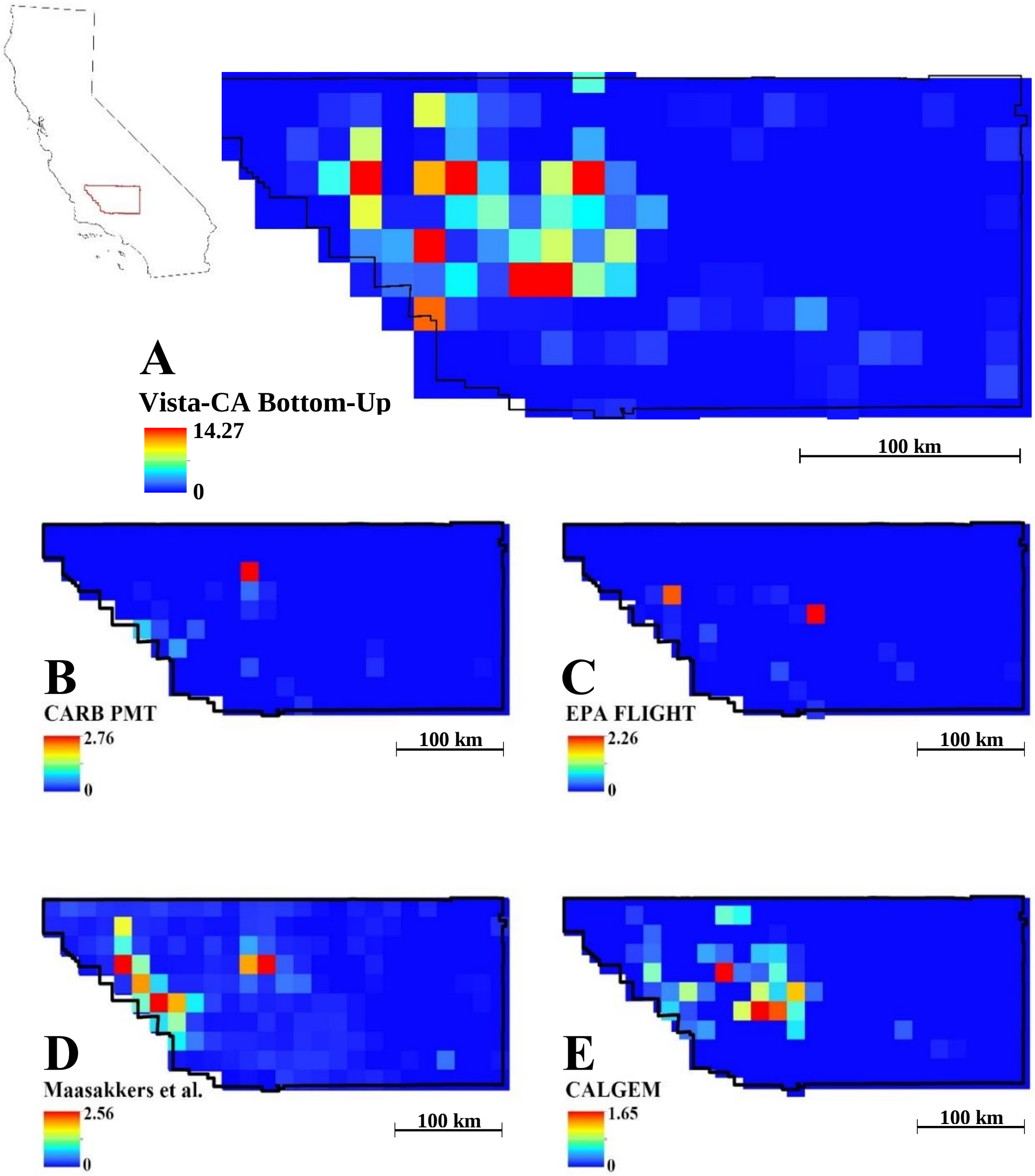


Figure 3.1 Kern County A) Vista-CA Bottom-Up CH_4 10 km by 10 km gridded data, B) CARB PMT facility CH_4 inventory, C) EPA FLIGHT facility CH_4 inventory, D) Maasakkers et al. 10 km x 10 km CH_4 gridded emissions data, and E) CALGEM 10 km x 10 km CH_4 gridded emissions data for Kern County. All data is presented in Gg CH_4 a^{-1}. (CARB PMT and EPA FLIGHT location data was converted into a 10 km x 10 km gridded format for comparison).

Figure 3.2 Vista-CA Kern County Oil and Gas dataset.

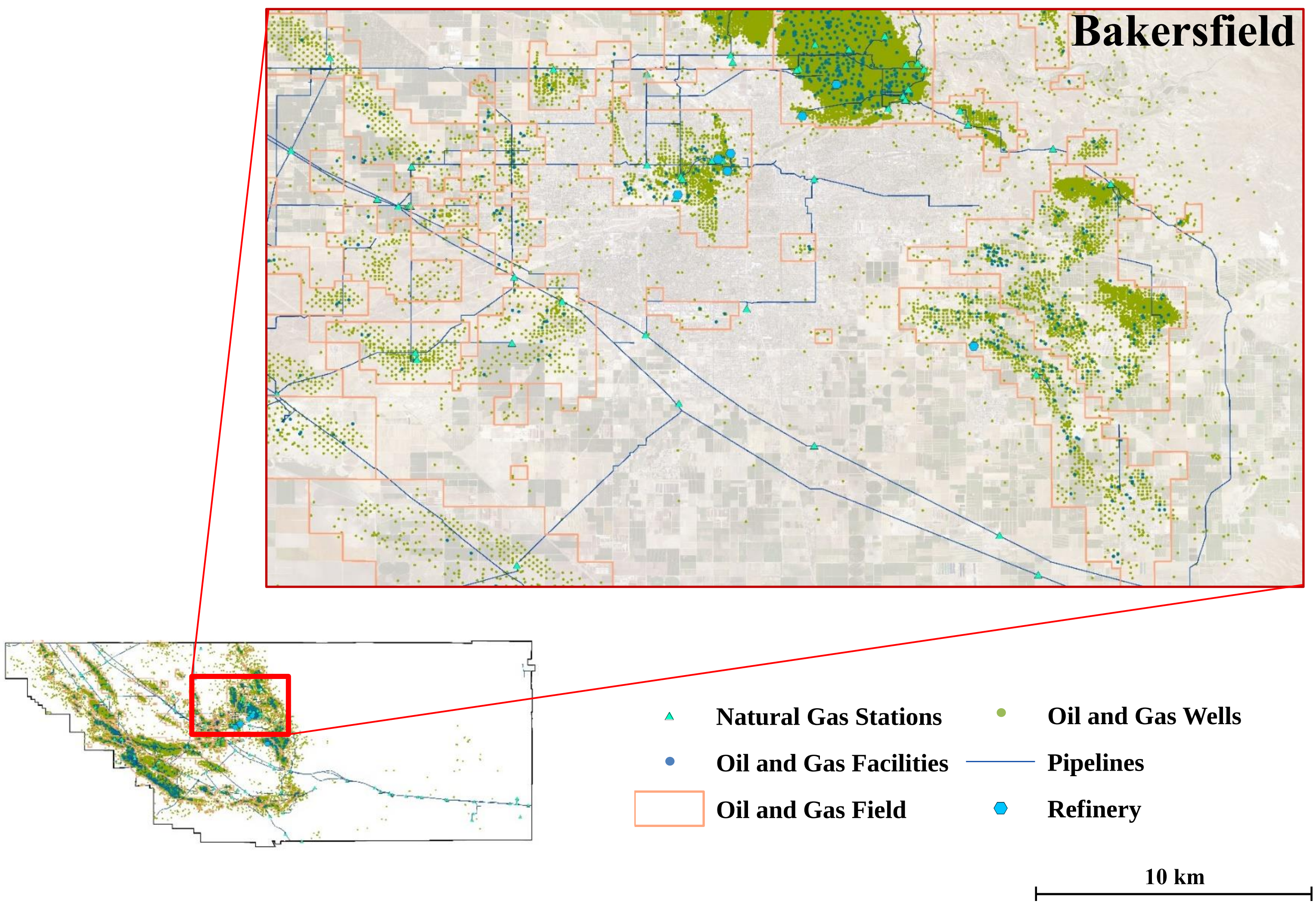


Figure 3.3 Vista-CA Kern County Bottom-Up CH_4 emissions dataset (1 km x 1 km grid).

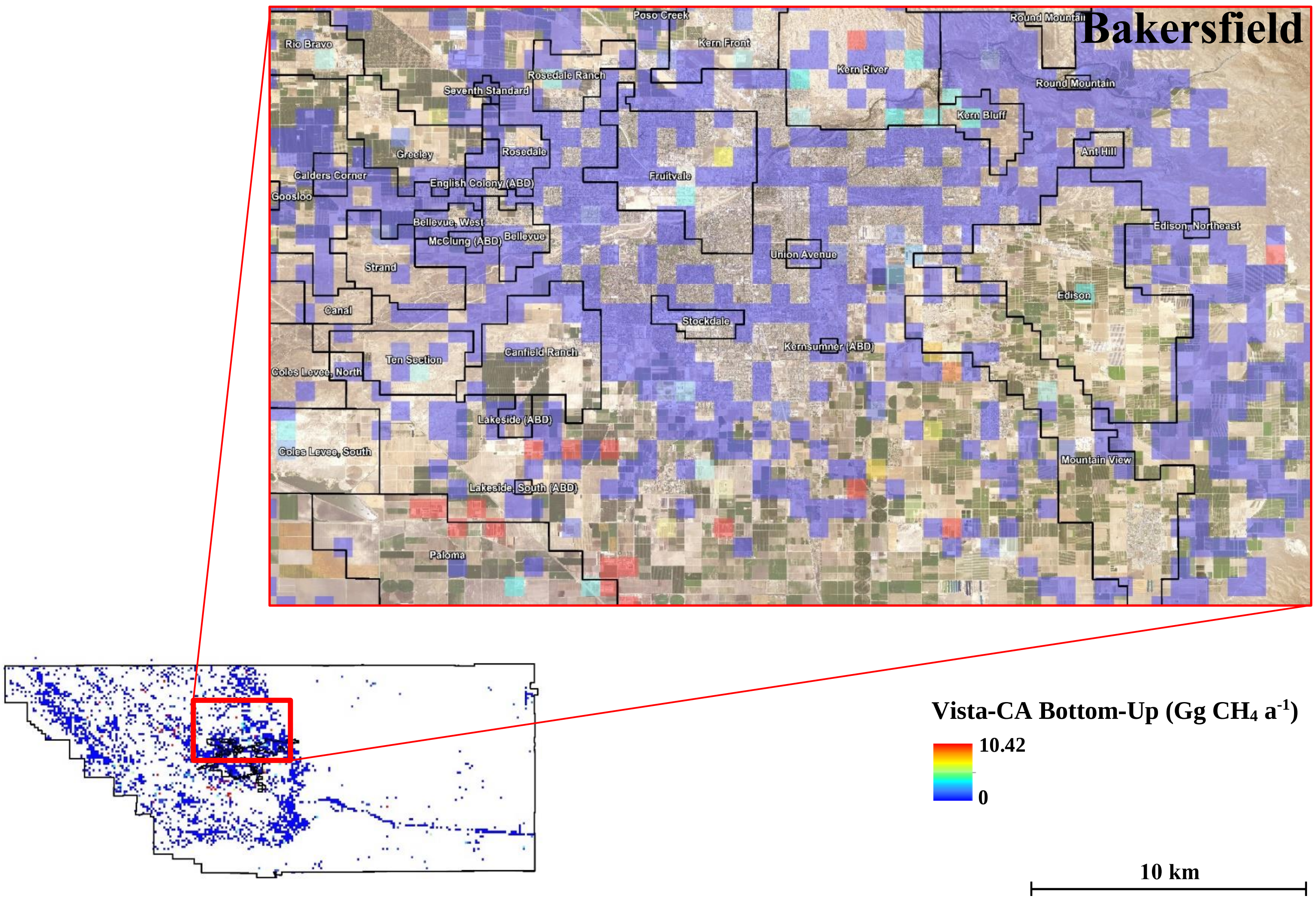
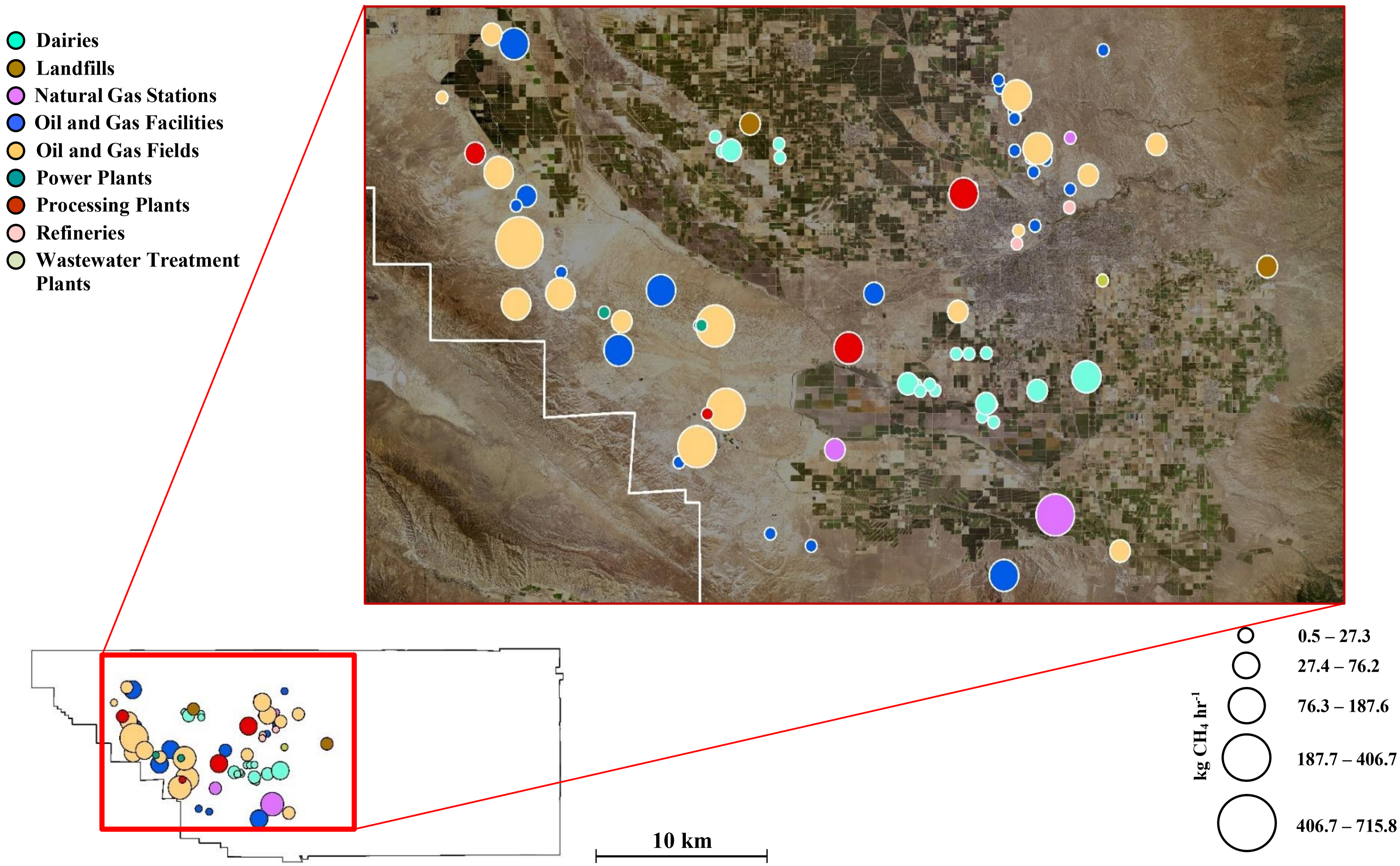


Figure 3.4. AVIRIS-NG Source Data product. This map displays 79 attributed Vista-CA facilities where 631 anomalous plumes were detected by AVIRIS-NG between 2016 – 2018, ranging from 0.5 – 715.8 kg CH_4 hr^{-1}. The colors represent the type of facility and the size of the circle represents the intensity of CH_4 emissions.



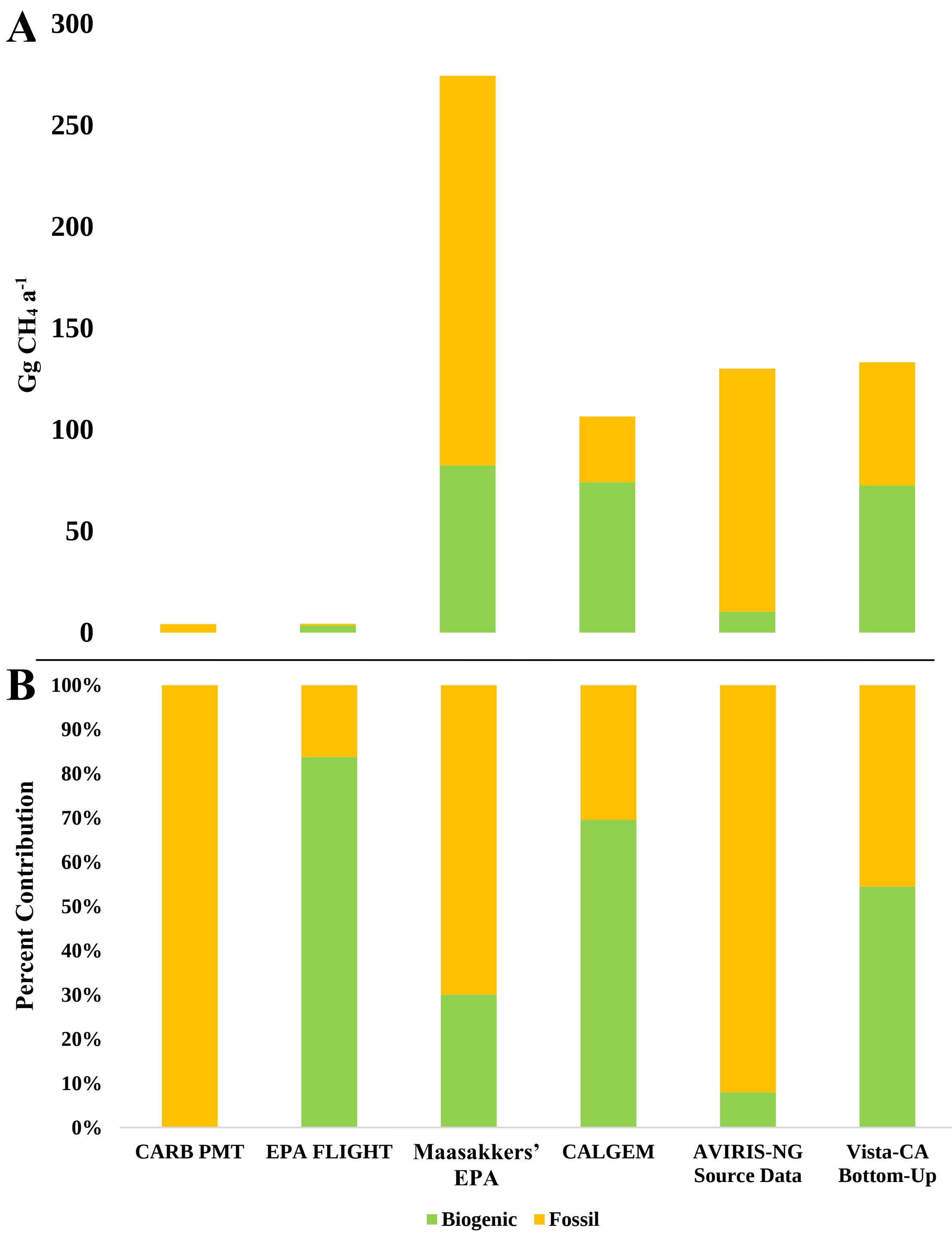


Figure 3.5 A) Magnitude of total estimated emissions (top) and B) Proportions by percentage of total emissions (bottom) classified into biogenic (green) and fossil (yellow) emission type for each dataset (Gg CH_4 a^{-1}).

Figure 3.6. Source apportionment for all datasets organized into IPCC Level 3 categories.

V. Supplementary

S3.1. Introduction

This document outlines the specific steps and procedures taken to develop and construct the data, methods, and analysis summarized in the main paper. The Vista-CA dataset is available at https://doi.org/10.3334/ORNLDAAC/1726. Attribution products from the Geospatial Source Attribution Automated Model (GSAAM) are available by request (https://doi.org/10.1088/1748-9326/ab9af8). The Next Generation Airborne Visible / Infrared Imaging Spectrometer (AVIRIS-NG) dataset is available at https://data.carbonmapper.org. Information on the California Greenhouse Gas Emissions Measurement (CALGEM) dataset can be found at http://calgem.lbl.gov/about.html and data is available by request. Information about the Maasakkers et al. 2016 dataset can be found at https://www.epa.gov/ghgemissions/gridded-2012-methane-emissions#data. EPA Facility Level Information on GreenHouse gases Tool (FLIGHT) dataset and information can be found at https://ghgdata.epa.gov/ghgp/main.do#. California Air Resource Board Pollution Mapping Tool (CARB PMT) dataset and information can be found at https://ww3.arb.ca.gov/ei/tools/pollution_map/. All emission estimates were converted into Gg CH_4 a^{-1}.

S3.2. Vista-CA Bottom-Up Data Development

In the following sections we provide a sectoral breakdown for the methods used to estimate CH_4 emissions. In order to generate the gridded product, we first converted all polygons and polylines to points. We appended emissions to each point accordingly and merged all points to one feature dataset. We converted the points to raster format dataset and summed the CH_4 emission values within each 1 km x 1 km and 10 km x 10 km grid cell. Finally, we standardized the gridded product.

We calculated CH_4 emissions at the facility level for each emission source feature in Kern County as included in Vista-CA (Rafiq et al. 2020). Detailed emission estimates for dairy farms, feed lots, and digesters were drawn from Marklein et al. (2021), which uses emission factors from CARB along with facility level herd data and regional manure management information. Oil and gas related emissions including fields, facilities, and wells were taken from the Vafi et al. (in review 2021), which uses the Oil Production Greenhouse Gas Emissions Estimator (OPGEE) model to calculate CH_4 emissions on site from within oil fields. Power plant emissions utilizing Intergovernmental Panel on Climate Change bottom-up modeling methods using the base equation: $E_{CH4} = Sum\ of\ types\ (Activity_{type} \times EF_{type})$, were taken from Rafiq et al. (2022). When more limited information was available which was the case for landfills, composting sites, and wastewater treatment plants, emissions were spatially disaggregated from previously calculated estimates in the U.S. EPA/Maasakkers et al. 2016 dataset. Emissions for natural gas compressor stations, natural gas processing plants, and transmissions pipelines were calculated using emission factors from the 2020 U.S. EPA Greenhouse Gas Inventory (GHGI) (GHGI 2020). Finally, reported emissions from CARB and EPA were used to allocate to Kern County refineries.

S3.2.1 Composting Sites

Vista-CA contains records for 18 composting sites for Kern County. Vista-CA utilized records from the California Department of Resources Recycling and Recovery (CalRecycle). Downscaled CH_4 emissions of composting sites do not currently exist in publicly available data. Instead, we used data from the U.S. EPA/Maasakkers et al. 2016 where they calculated state-level emissions from composting using the tonnage of municipal solid waste composted or from the correlated tonnage recycled. Emissions were assigned to locations from the US Composting Council, BioCycle composter database, and composting entries in the Facility Registry Service (FRS). If there were fewer than three facilities found in a given state, they allocated them based on the gridded population instead (Maasakkers et al. 2016). We utilized this emissions allocation for composting sites totaling 0.76 Gg CH_4 a^{-1} and divided it evenly across the 18 composting sites that resulted in a per facility estimate of 42,000 kg CH_4 a^{-1}.

S3.2.2 Natural Gas (Compressor) Stations

Vista-CA contains records for 142 stations for Kern County. We utilized 2020 U.S. EPA Greenhouse Gas Inventory (GHGI) for natural gas emissions to obtain emission factors in kg per unit activity based on the type of compressor station (GHGI 2020). EPA developed these emission factors through either random or stratified sampling and then summing these samples and dividing by the number of total sources sampled. Further information can be found at https://www.epa.gov/natural-gas-star-program/methane-emissions-natural-gas-industry. Specifically for natural gas stations, they took the number of stations from the Greenhouse Gas Reporting Program (GHGRP) and multiplied it by a scaling factor of 3.52 (from Zimmerie et al. 2015). Natural gas station type was stored in the Vista-CA metadata. The following breaks down the type of stations that were in the dataset and the respective emissions: 22x COMP: compressor station, 310,708.7 kg/station a^{-1}, 6x DEHY: dehydration station, 50x MTR: metering station, 2,142.7 kg/station a^{-1}, 9x

ODOR: odor stations, 17x PLS: pressure limiting stations, 22x REG: regulation stations, 15x STR: storage stations, 369,797.4 kg/station a^{-1}, and 1x TAP stations. Emissions from dehydration stations, odor stations, pressure limiting stations, regulations stations, storage stations, and TAP stations are considered negligible. Emissions were converted into kg CH_4 a^{-1}.

S3.2.3 Dairies

Vista-CA contains records for 53 dairies for Kern County. We utilized emission estimates for Kern County dairies from Marklein et al.'s work (Marklein et al. 2021). Marklein et al. developed a spatially explicit database of dairies in California using information from operating permits and California-specific reports detailing herd demographics and manure management at the facility scale. They estimated manure management and enteric fermentation CH_4 emissions utilizing previously published bottom-up approaches and a new farm-specific calculation. Marklein et al.'s data contained emissions for 44 of the 53 dairies in Kern County. The remaining 9 dairies were not considered significant emitters after additional correspondence with the author of Marklein et al.'s work and final emission estimates were converted into kg CH_4 a^{-1}.

S3.2.4 Digesters

Vista-CA contains records for 5 digesters in Kern County. We utilized emission estimates for Kern County dairies from Marklein et al.'s work (Marklein et al. 2021). From the breakdown below we used the emissions from the "digester_ch4" category.

Vista_ID	post_digester_total_manure	digester_ch4	total_enteric	total_ch4_postdigester	predigester_total_ch4
DAF000177	0.803584091	0.293611428	1.576858249	2.38044234	3.261276624
DAF000178	0.091628553	0.050888816	0.226591068	0.318219621	0.470886069
DAF001720	0.489163453	0.106964448	1.064318054	1.553481507	1.87437485
DAF000259	0.64722218	0.236480327	1.294048798	1.941270978	2.650711959
DAF000983	0.435801798	0.159232107	0.823133992	1.25893579	1.736632112

These emissions were calculated using the following description from Marklein et al.: "We determined the 106 dairies that have installed or are planning to install anaerobic digesters from reports from the CDFA Dairy Digester Reports in 2017–2019 (CDFA, 2020b). We used our database to estimate the effects of anaerobic digesters on CH_4 emissions from these 100 dairies in the Central Valley. We assumed a 75 % efficiency of CH_4 capture in anaerobic digesters (Charrier, 2016; US EPA, 2017)." "Post_digester_total_manure" field refers to the total manure CH_4 emissions when you consider that the digester is in use, "Total_CH_4_postdigester" refers to the total manure + enteric CH_4 emissions when you consider that the digester is in use, and "Predigester_total_CH_4" refers to the total manure and enteric CH_4 emissions if the facility did not have a digester. Emissions were converted into kg CH_4 a^{-1}.

S3.2.5 Feed Lots

Vista-CA contains records for 2 feed lots in Kern County. We used the total number of cows in Kern County from the United States Department of Agriculture's National Agricultural Statistics Service Information (USDA NASS) livestock census from 2017 and assumed the population remained constant in 2018 because the emission factors remained the same. The USDA NASS reports 41,537 cows, which we divided evenly between the 2 feedlots in Kern County, for a total of 20,768.5 cows per farm. We used emissions estimates from the 2018 CARB's inventory to determine the CH_4 emissions per head of feedlot animals. The ARB reports 715.97 tons of CH_4 emitted by 330,227 head of feedlot steers 500+lbs and 383.75 tons CH_4 produced by 171,796 head feedlot heifers that were 500+ lbs. This equates to 2.171052 kg/head steers 500+lbs and 2.233754 kg /head of feedlot heifer 500+ lbs. By dividing the populations of steers and heifers by the total cattle heads, we determined that 65% of the animals that were 500+lbs were steer and 35% were heifers. We took a weighted average of the total CH_4 per head to estimate the average CH_4 emissions/head to be 2.192998 kg/head (Marklein et al. 2021). We multiplied the population on each farm by the emissions factor Alison calculated below, 2.19 kg CH_4 per head, to get 45,483.015 kg CH_4 per farm a^{-1}.

S3.2.6 Landfills

Vista-CA contains records for 107 landfills in Kern County. We obtained CARB 2014 emissions data for 17 landfills in Kern County. We utilized U.S. EPA emissions for county wide data for municipal landfills for the remaining 90 landfills without emissions through budget allocation (Maasakkers et al. 2016). U.S. EPA allocated emissions from landfills based on a combination of data from the GHGRP (1,231 municipal landfills and 175 industrial landfills), the Landfill Methane Outreach Program (LMOP) (municipal landfill only), and the FRS. GHGRP landfills were assigned their reported emissions and they utilized waste-in-place combined with a decay factor for the other landfills. For the unassigned landfills they assigned the median emissions from the landfills for which information was available (Maasakkers et al. 2016). County wide data totaled at 13.51 Gg CH_4 a^{-1}. We subtracted the CARB 2014 emissions which totaled 5.012 Gg CH_4 a^{-1}. We allocated the remaining 5.69 Gg CH_4 a^{-1} evenly for the remaining 90 facilities at 63,314 kg CH_4 a^{-1} per facility.

S3.2.9 Oil and Gas Fields

Vista-CA contains records for 100 oil and gas fields in Kern County and also contains records for 1,888 oil and gas facilities as well as 142,534 oil and gas wells. We obtained CH_4 emissions from Oil Production Greenhouse Gas Emissions Estimator (OPGEE) model which calculates CH_4 emissions from oil and gas fields and also incorporates emissions from oil and gas wells as well as oil and gas facilities found within the bounds of these oil and gas fields (Note: while OPGEE only includes emissions from power plants associated with electricity consumption to produce crude oil, Vafi et al. 2021 calculated these separately for their onsite emissions since the model includes only LCA emissions). OPGEE is a life cycle assessment model that includes offsite emissions and borrows emissions factors from CARB and the American Petroleum Institute (API) which was reorganized to report only onsite emissions by Vafi et al. 2021. Vafi et al. utilized OPGEE 2.0 to produce emissions for 145 oil and gas fields in California which include all 100 oil and gas fields in Kern County (Vafi et al. 2021). These also incorporate 1,878 oil and gas facilities and 139,235 oil and gas wells. The remaining oil and gas wells and oil and gas facilities from Vista-CA are calculated separately and are discussed in their respective sections. Emissions were converted into kg CH_4 a^{-1}.

S3.2.8 Oil and Gas Facilities

Vista-CA contains records for 1,888 oil and gas facilities in Kern County. We calculated emissions for all 1,888 oil and gas facilities and then at the end removed 1,878 oil and gas facilities as they were already included in the OPGEE modeled emissions for oil and gas fields. We utilized county-wide U.S. EPA/Maasakkers et al. 2016 emissions for

Petroleum facilities at 154.44 Gg CH_4 a^{-1}, Natural Gas Distribution facilities 0.84 Gg CH_4 a^{-1}, Natural Gas Processing facilities 15.31 Gg CH_4 a^{-1}, Natural Gas Production facilities 19.83 Gg CH_4 a^{-1}, and Natural Gas Transmission 0.73 Gg CH_4 a^{-1} (Maasakkers et al. 2016). Within the Vista-CA Oil and Gas Facilities data, we utilized the metadata to identify and remove facilities based on status, specifically 81 no status and 392 idle/out of service/removed were removed from the list leaving 1,415 active and deserted oil and gas facilities. 369 of those 1,415 facilities were combo gas/oil, gas facilities and we utilized the U.S. EPA/Maasakkers et al. 2016 processing and production number of 35.14 Gg CH_4 a^{-1} and divided by 369 to obtain 95,230.35 kg CH_4 a^{-1} per facility. For the remaining 1,026 facilities, we utilized the Petroleum number of 154.44 Gg CH_4 a^{-1} and divided among the 1,026 facilities to obtain an emission estimate of 150,526.316 kg CH_4 a^{-1} per facility. The remaining 3 facilities tagged as injection, steam and treatment were negligible. The final dataset contains emissions for 10 oil and gas facilities independent of oil and gas fields and emissions were converted into kg CH_4 a^{-1}.

S3.2.9 Oil and Gas Wells

Vista-CA contains records for 142,534 oil and gas wells in Kern County. We calculated emissions for all 142,534 oil and gas wells and then at the end removed 139,235 wells as they were already included in the OPGEE modeled emissions for oil and gas fields. We obtained emission factors from the 2020 U.S. EPA GHGI for specific oil and gas wells (GHGI 2020). There are 64,258 oil and gas wells whose status are plugged or abandoned which are treated as negligible based on Fischer et al.'s work. Fischer et al. stated "The limited data presented [in their work] suggest that CH_4 emissions from abandoned/plugged

(AP) wells in California are negligible, at least for wells located primarily outside large active oil and gas fields. Excluding measurements from wells in the tar pit (which may have their own natural sources), the authors found no AP wells to be leaking above 1 g/hr." (Fischer et al. 2020). The remaining 78,276 oil and gas wells had statuses of active, buried, idle, new, and unknown. 15,824 of the 78,276 oil and gas wells were identified as directionally drilled and were assigned an emission factor of 51 kg CH_4 a^{-1}. The other 62,452 were non-directional drilling. 407 of the 62,452 oil and gas wells were assigned 89.4 kg CH_4 a^{-1} they were identified as the following types and were non-hydraulic fracturing: Air Injector/Oil and Gas, Dry Gas/Oil and Gas, Gas Disposal, Gas Disposal/Oil and Gas/Pressure Maintenance/Cyclic Steam, Observation/Oil and Gas/Cyclic Steam, Oil and Gas, Oil and Gas/Pressure Maintenance, Oil and Gas/Cyclic Steam, Oil and Gas/Cyclic Steam/Steam Flood, Oil and Gas/Cyclic Steam/Steam Flood/Water Disposal/Water Flood, Oil and Gas/Cyclic Steam/Steam Flood/Water Flood, Oil and Gas/Cyclic Steam/Water Disposal, Oil and Gas/Steam Flood, Oil and Gas/Steam Flood/Water Flood, Oil and Gas/Water Disposal, Oil and Gas/Water Disposal/Water Flood, Oil and Gas/Water Flood, Oil and Gas/Water Source. The following 62,045 oil and gas well types were also treated as negligible since they did not have anything related to natural gas: Air Injector, Observation/Steam Flood, Pressure Maintenance, Cyclic Steam, Cyclic Steam/Steam Flood, Cyclic Steam/Water Flood, Steam Flood, Steam Flood/Water Disposal, Steam Flood/Water Flood, Water Disposal, Water Disposal/Water Flood, Water Flood, Water Source. The final dataset contains emissions for 3,299 oil and gas wells and emissions independent of oil and gas fields were converted into kg CH_4 a^{-1}.

S3.2.10 Natural Gas Processing Plants

Vista-CA contains record for 9 natural gas processing plant facilities in Kern County. We utilized 2020 U.S. EPA GHGI for natural gas emissions to obtain emission factors in kg per unit activity based on the type of compressor station (GHGI 2020). EPA developed these emission factors through either random or stratified sampling and then summing these samples and dividing by the number of total sources sampled. Further information can be found at https://www.epa.gov/natural-gas-star-program/methane-emissions-natural-gas-industry. Specifically, for natural gas processing plants, they recorded the number of facilities from the Oil and Gas Journal "Worldwide Gas Processing" survey for the years 1997-2014. This survey contained information on gas capacity, gas throughput, and total natural gas produced per day for 29 facilities in California. We utilized the calculated emission factors from the GHGI report for natural gas processing plants at 212,382.3 kg CH_4 a^{-1} per facility.

S3.2.11 Pipelines

Vista-CA contains records for 3,696 transmission pipeline segments that span 2173.13 km total. We obtained emission factors from the 2020 EPA Greenhouse Gas Inventory (GHGI) specifically for pipeline leaks under “transmission and storage” calculated at a rate of 10.9 kg/mile (GHGI 2020). EPA developed these emission factors through either random or stratified sampling and then summing these samples and dividing by the number of total sources sampled. Further information can be found at https://www.epa.gov/natural-gas-star-program/methane-emissions-natural-gas-industry.

EPA utilized the total transmission pipeline mileage for a given year using data from the

Pipeline and Hazardous Materials Safety Administration (PHMSA). We obtained a total emission estimate of 14,718.53 kg of CH_4 a^{-1} for the pipeline sector. Emissions were converted into kg CH_4 a^{-1}.

S3.2.12 Power Plants

Vista-CA contains records for 43 power plants total in Kern County. 35 of these plants contained records from Rafiq et al. 2022. Rafiq et al. utilized IPCC tier 2 methods for stationary combustion to model power plant emissions which multiplies country-specific emission factors by fuel-type with total fuel combusted. Specifically, they utilized annual fuel consumption metrics from the U.S. Energy Information Administration (EIA), emission factors from the EPA's Electronic Code of Federal Regulations (ECFR) and plant type from Vista-CA to establish bottom-up estimates of power plants in California. Three additional plants from the CARB PMT matched with the remaining 8 but contained no emissions or were not significant. We were unable to match the other 5 power plants with CARB or EPA data. Emissions were converted into kg CH_4 a^{-1}.

S3.2.13 Refineries

Vista-CA contains records for 7 refineries total in Kern County. Four of these refineries contained emission records from CARB and EPA. Both CARB and EPA contain reported emissions from these refineries. The remaining 3 refineries were individual facilities that shared names and possible infrastructure with the other four existing refineries (from aerial imagery analysis). These 3 refineries were Alon Bakersfield Refining 3, San Joaquin Refining Co, and Alon Bakersfield Operating Inc. Emissions were converted into kg CH_4 a^{-1}.

S3.2.14 Wastewater Treatment Plants

Vista-CA contained records for 7 wastewater treatment plants in Kern County. We utilized U.S. EPA emissions for wastewater treatment plants. Emissions from wastewater treatment were reported as municipal or industrial in the GHGI and facilities that reported to the GHGRP accounted for the majority of the national industrial wastewater treatment emissions. To allocate the remaining industrial emissions as well as the municipal emissions, they used facility-level wastewater flow data from the Clean Watersheds Needs Survey (Maasakkers et al. 2016). We utilized U.S. EPA emissions for county wide data for domestic and industrial wastewater treatment plants at 1.08 Gg CH_4 a^{-1}. We allocated the emissions evenly for the 7 facilities at 150,000 kg CH_4 a^{-1} per facility.

S3.2.15 Negligible Emissions

The two Vista-CA sectors that were considered negligible for CH_4 emissions were storage fields and natural gas fueling stations. There were no storage fields located within Kern County. Vista-CA contains records for 5 natural gas fueling stations. Fischer et al. determined that fugitive emissions from natural gas stations were negligible (Fischer et al. 2017).

S3. AVIRIS-NG Source Data Development

In order to account for gaps in observation from AVIRIS-NG into our emissions calculation, we developed data that identifies all plumes flown and all flightlines that flew over those plumes and for each plume whether that flightline that flew over it, observed that given plume or did not observe that given plume.

We used the AVIRIS-NG 2016, 2017, and 2018 plume data that contains 1,704 plumes across California. We geocoded source locations of each plume using the source latitude and source longitude in ArcGIS into point vector locations. We also used line vector data for all 1,943 flightlines flown in 2016, 2017, and 2018 AVIRIS-NG campaigns. We generated 1,800 m wide buffer polygon vectors from the line vectors of each flightline to simulate projected flight path or swath width of the AVIRIS instrument. Next, we took both the point vector data and the polygon vector data and utilized the intersect tool in ArcGIS. This intersect tool allowed us to compute thousands of geometric intersections between the point and polygon vector data. Any overlap between point and polygon vectors indicates that all possible flightpaths over a given plume location to identify how many times a plume was flown over and how many times it was actually observed by AVIRIS-NG. In total, this resulted in 28,989 plume-to-flightpath spatial connections. Consequently, we wanted to identify which plumes were actually flown by the flightline that was spatially associated with it and which plumes were not. In order to do this, we extracted out the tabular file from our previous 28,989 plume-to-flightpath data for further development. We added a new field called “Plume_Detected?” that evaluates whether the “linename” field that comes from the plumes data matches the “LineName” field from the flightlines dataset. Essentially, if these two fields matched that meant that a given plume observed and spatially associated to a specific flightline was indeed observed by that specific flight line, which in this case meant it was marked with a “yes”. Conversely, if those two fields did not match, meaning a given plume observed and spatially associated with a given flightline did not match with the flightline that originally observed with that plume, then we take this as the flightline flew over that given location but did not observe that given plume and mark it with a “no” in the “Plume_Detected?” field. The total number of “yes”s should exactly equal 1,704 since that’s how many total plumes we have in the AVIRIS-NG dataset. Our spatial and tabular methods resulted in 1,579 “yes”s which still doesn’t account for the missing 125 plumes observed. This is because some of the flightline buffers we generated using the centerline and swath of 1,800 m of each flightline are assumed to be uniform and ideal and are not completely based on actual paths flown. Therefore, we had to manually locate the missing 125 “yes” plumes. These 125 additional “yes” plumes were identified by comparing the 1,579 “yes” plumes to the original 1,704 plumes using the “Candidate_ID” field and those that didn’t match were appended. In this case, these 125 additional “yes” plumes were added to the existing 1,579 “yes” plumes bringing the total to 1,704 and also brought the sum total of 28,989 to 29,114 total plume-to-flightpath spatial connections.

To obtain the final product as we first defined, we needed to remove redundant plume-to-flightpath connections that are identified as “no” under the “Plume_Detected?” field. The term “redundant connections” refers to all possible spatial connections made for plumes of a given source that were already labeled by the “Plume_Detected?” field as “no” to begin with. To do this, we created a new temporary identification for each plume-to-flightpath connection by concatenating the plume “Source_ID” field and flightpath “LineName” field in the dataset in order to systematically isolate the redundant spatial connections. We identified and removed duplicates based on this temporary identification field and only kept the connections that were labeled as unique by this temporary identification field. This process reduced our total plume-to-flightpath spatial connections to 9,473. Additionally, we identified 9 spatial connections that were removed by our redundancy process that actually needed to be added to our total connections list. These 9 connections were identified because in the original AVIRIS-NG plume dataset, there are 9 repeating "Candidate_ID"s with different emissions estimates and thus they were not detected by our methods early on until we conducted checks at the very end for consistency. After verifying and adding in these 9 spatial connections, the source data resulted in a total of 9,482 plume-to-flightpath spatial connections stemming from 795 unique sources as identified by AVIRIS-NG. Lastly, we ran the finalized source data product through Vista-CA GSAAM to obtain Vista-CA facility and sub-facility spatial attributions, along with explicit identification of Oil and Gas Field spatial attributions for all 9,482 plume-to-flightpath spatial connections. We have taken the total number of times a given facility in each study area was flown by AVIRIS-NG and recorded the number of times it was observed to be emitting and observed to not be emitting.

Tables S3.1 and S3.2 show facility breakdowns by sector and region. They also highlight emissions estimates and survey coverages. This data in both geospatial and tabular formats can be extracted from the Vista-CA and AVIRIS-NG datasets listed above.

S4. Ancillary Emissions Data Calculation Methods

We standardized both the CALGEM and Maasakkers et al. dataset from netcdf format to a gridded raster format. The following sections breakdown the methods employed for emission estimation for each sector in each dataset.

S4.1 CALGEM Sectoral Emissions Estimate Methods

Jeong et al. 2016 developed CALGEM prior emission distributions and scaled them to match 2012 California Air Resources Board (CARB) state totals for anthropogenic emission sectors (CARB, 2014), with small (<50 Gg CH_4 a^{-1}) adjustments for some regions and sectors. 2012 CARB data is based off of methods developed in the 2006 Intergovernmental Panel on Climate Change (IPCC) report. Other sectors used data from ancillary state government sources. Raw data was converted into Gg CH_4 a^{-1} for this work.

S4.1.1 Livestock/Non-livestock: Jeong et al. 2016 revised the spatial distribution of the dairy livestock emissions by incorporating the 2012 county-level dairy statistics from U.S. Department of Agriculture (http://www.nass.usda.gov/Statistics_by_State/California/Publications/County_ Estimates/) to the spatial distribution from Jeong et al. [2013]. They also used a map of dairy livestock density supplied by the California Department of Water Resources and scaled to annual CH_4 emissions assuming a constant emission factor of 0.39 kg C/cow/d from Salas et al. 2009.

S4.1.2 Natural Gas: Jeong et al. 2016 used information from the California Department of Conservation (CDC, http://www.consrv.ca.gov/dog/Pages/statistics.aspx) to generate an emissions map of Natural Gas wells in California. CH_4 emissions from gas wells were estimated using gas production information from California Department of Conservation [2009]. They also assumed a leakage rate of 1% related to gas production and transmission/storage processes in the gas fields. Some CH_4 emissions were estimated from California mandatory reports on oil and gas and the remainder of natural gas emissions was apportioned by population density in California using 4 km population maps available from the Socioeconomic Data and Applications Center (Center for International Earth Science Information Network (CIESIN) and Centro Internacional de Agricultura Tropical (CIAT), Gridded population of the world, version 3, population density grid, 2005, http://sedac.ciesin.columbia.edu/gpw).

S4.1.3 Wastewater Treatment: Jeong et al. 2016 used the 2012 CARB state totals for anthropogenic emissions for this sector [CARB, 2014]. The 2012 CARB wastewater treatment plant emissions data is based on the 2006 IPCC report. Description below shows the Tier 1 equations for calculating domestic and industrial wastewater treatment plant emissions are found here: 2006 IPCC Report Volume 5 Ch. 6, Pg. 6.11 and 6.20.

(https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/5_Volume5/V5_6_Ch6_Wastewater.pdf)

The general equation to estimate CH_4 emissions from domestic wastewater is as follows:

EQUATION 6.1
TOTAL CH_4 EMISSIONS FROM DOMESTIC WASTEWATER

$$CH_4\ Emissions = \left[\sum_{i,j}\left(U_i \bullet T_{i,j} \bullet EF_j\right)\right]\left(TOW - S\right) - R$$

Where:

CH_4 Emissions = CH_4 emissions in inventory year, kg CH_4/yr

TOW = total organics in wastewater in inventory year, kg BOD/yr

S = organic component removed as sludge in inventory year, kg BOD/yr

U_i = fraction of population in income group *i* in inventory year, See Table 6.5.

$T_{i,j}$ = degree of utilisation of treatment/discharge pathway or system, *j*, for each income group fraction *i* in inventory year, See Table 6.5.

i = income group: rural, urban high income and urban low income

j = each treatment/discharge pathway or system

EF_j = emission factor, kg CH_4 / kg BOD

R = amount of CH_4 recovered in inventory year, kg CH_4/yr

S4.1.4 Petroleum Refining and Mobile Sources: Jeong et al. 2016 used the 2012 CARB state totals for anthropogenic emission for this sector [CARB, 2014]. The 2012 CARB refinery and mobile sources emissions data is based on the 2006 IPCC report. Tier 1 equations for stationary combustion emissions are found here: 2006 IPCC Report Volume 2 Ch. 2, Pg. 2.11. (https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/2_Volume2/V2_2_Ch2_Stationary_Combustion.pdf)

Applying a Tier 1 emission estimate requires the following for each source category and fuel:

- Data on the amount of fuel combusted in the source category
- A default emission factor

Emission factors come from the default values provided together with associated uncertainty range in Section 2.3.2.1. The following equation is used:

EQUATION 2.1
GREENHOUSE GAS EMISSIONS FROM STATIONARY COMBUSTION

$$Emissions_{GHG,\,fuel} = Fuel\ \ Consumption_{fuel} \bullet Emission\ \ Factor_{GHG,\,fuel}$$

Where:

$Emissions_{GHG,fuel}$ = emissions of a given GHG by type of fuel (kg GHG)

$Fuel\ Consumption_{fuel}$ = amount of fuel combusted (TJ)

$Emission\ Factor_{GHG,fuel}$ = default emission factor of a given GHG by type of fuel (kg gas/TJ). For CO_2, it includes the carbon oxidation factor, assumed to be 1.

Tier 1 equations for calculating mobile combustion CH_4 emissions can be found here: 2006 IPCC Report Volume 2 Ch. 3, Pg. 3.13. (https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/2_Volume2/V2_3_Ch3_Mobile_Combustion.pdf).

The equation for the Tier 1 method for estimating CH_4 and N_2O from road vehicles may be expressed as:

EQUATION 3.2.3
TIER 1 EMISSIONS OF CH_4 AND N_2O

$$Emission = \sum_{a} [Fuel_a \bullet EF_a]$$

Where:

Emissions = emission in kg

EF_a = emission factor (kg/TJ)

$Fuel_a$ = fuel consumed, (TJ) (as represented by fuel sold)

a = fuel type a (e.g., diesel, gasoline, natural gas, LPG)

Equation 3.2.3 for the Tier 1 method implies the following steps:

- Step 1: Determine the amount of fuel consumed by fuel type for road transportation using national data or, as an alternative, IEA or UN international data sources (all values should be reported in terajoules).
- Step 2: For each fuel type, multiply the amount of fuel consumed by the appropriate CH_4 and N_2O default emission factors. Default emission factors may be found in the next Section 3.2.1.2 (Emission Factors).
- Step 3: Emissions of each pollutant are summed across all fuel types.

S4.1.5 Landfills: Jeong et al. 2016 used the 2012 CARB state totals for anthropogenic emission for this sector [CARB, 2014]. The 2012 CARB landfills emissions data is based on the 2006 IPCC report. Tier 1 equations for calculating landfill CH_4 emissions are found here: 2006 IPCC Report Volume 5 Ch. 3, Pg. 3.8. (https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/5_Volume5/V5_3_Ch3_SWDS.pdf)

The CH_4 emissions from solid waste disposal for a single year can be estimated using Equations 3.1. CH_4 is generated as a result of degradation of organic material under anaerobic conditions. Part of the CH_4 generated is oxidised in the cover of the SWDS, or can be recovered for energy or flaring. The CH_4 actually emitted from the SWDS will hence be smaller than the amount generated.

EQUATION 3.1
CH_4 EMISSION FROM SWDS

$$CH_4\ Emissions = \left[\sum_x CH_4\ generated_{x,T} - R_T\right] \bullet (1 - OX_T)$$

Where:

CH_4 Emissions = CH_4 emitted in year T, Gg

T = inventory year

x = waste category or type/material

R_T = recovered CH_4 in year T, Gg

OX_T = oxidation factor in year T, (fraction)

The CH_4 recovered must be subtracted from the amount CH_4 generated. Only the fraction of CH_4 that is not recovered will be subject to oxidation in the SWDS cover layer.

S4.2 Maasakkers et al. 2016 Sectoral Emissions Estimate Methods

Maasakkers et al. 2016 used the 2012 emissions from the 2014 EPA Greenhouse Gas Inventory (GHGI) published in 2016, which includes detailed descriptions of the national emission calculation methods. GHGI derives most of their methods from the 2006 IPCC Report. Raw data was converted into Gg CH_4 a^{-1} for this work.

S4.2.1 Landfills: Landfill emissions were estimated as the CH_4 produced from MSW landfills, plus the CH_4 produced by industrial waste landfills, minus the CH_4 recovered and combusted from MSW landfills, minus the CH_4 oxidized before being released into the atmosphere; The methodology for estimating CH_4 emissions from landfills is based on the first order decay (FOD) model described by the 2006 IPCC Guidelines. Tier 1 equations for calculating landfill CH_4 emissions are found here: 2006 IPCC Report Volume 5 Ch. 3, Pg. 3.8. (https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/5_Volume5/V5_3_Ch3_SWDS.pdf) (See S4.1.5)

S4.2.2 Enteric Fermentation: Maasakkers et al. 2016 estimated CH_4 emissions from enteric fermentation by using the following method: the population was divided into state, age, sub-type (i.e., dairy cows and replacements, beef cows and replacements, heifer and steer stockers, heifers and steers in feedlots, bulls, beef calves 4 to 6 months, and dairy calves 4 to 6 months), and production (i.e., pregnant, lactating) groupings to more fully capture differences in CH_4 emissions from these animal types. They simulated the age and weight structure of each sub-type on a monthly basis in order to reflect the fluctuations that occur throughout the year. Cattle diet characteristics were used in conjunction with Tier 2 equations from 2006 IPCC report to produce CH_4 emission factors for the following cattle types: dairy cows, beef cows, dairy replacements, beef replacements, steer stockers, heifer stockers, steer feedlot animals, heifer feedlot animals, bulls, and calves. Information on tier 2 equations for calculating CH_4 emissions are found here: 2006 IPCC Report Volume 4 Ch. 10, Pg. 10.10. (https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/4_Volume4/V4_10_Ch10_Livestock.pdf) Finally, they estimated emissions from cattle using monthly population data and multiplying it by the calculated emission factor for each cattle type.

S4.2.3 Manure Management: Maasakkers et al. 2016 estimated CH_4 emissions from manure management by using the following method: the animal population data (by animal type and state): Typical animal mass (TAM) data (by animal type); Portion of manure managed in each WMS, by state and animal type; Volatile solids (VS) production rate (by animal type and state or United States); CH_4 producing potential (B_o) of the volatile solids (by animal type); and CH_4 conversion factors (MCF), the extent to which the CH_4 producing potential is realized for each type of WMS (by state and manure management system, including the impacts of any biogas collection efforts).

S4.2.4 Wastewater Treatment Plants: Maasakkers et al. 2016 estimated CH_4 emissions from publicly owned treatment works (POTWs) using EPA GHGI methodology. They began by multiplying the total Biochemical Oxygen Demand$_5$ (BOD_5) produced in the United States by the percent of wastewater treated centrally (about 80 percent), the relative percentage of wastewater treated by aerobic and anaerobic systems, the relative percentage of wastewater facilities with primary treatment, the percentage of B_{OD5} treated after

$$\textit{Emissions from Septic Systems} = A$$
$$= US_{POP} \times (\%\ onsite) \times (EF_{SEPTIC}) \times 1/10^9 \times Days$$

$$\textit{Emissions from Centrally Treated Aerobic Systems} = B$$
$$= [(\%\ collected) \times (total\ BOD_5\ produced) \times (\%\ aerobic) \times (\%\ aerobic\ w/out\ primary) + (\%\ collected) \times (total\ BOD_5\ produced) \times (\%\ aerobic) \times (\%\ aerobic\ w/primary) \times (1\text{-}\%\ BOD\ removed\ in\ prim.\ treat.)] \times (\%\ operations\ not\ well\ managed) \times (B_o) \times (MCF\text{-}aerobic\_not\_well\_man)$$

$$\textit{Emissions from Centrally Treated Anaerobic Systems} = C$$
$$= [(\%\ collected) \times (total\ BOD_5\ produced) \times (\%\ anaerobic) \times (\%\ anaerobic\ w/out\ primary) + (\%\ collected) \times (total\ BOD_5\ produced) \times (\%\ anaerobic) \times (\%\ anaerobic\ w/primary) \times (1\text{-}\%\ BOD\ removed\ in\ prim.\ treat.)] \times (B_o) \times (MCF\text{-}anaerobic)$$

$$\textit{Emissions from Anaerobic Digesters} = D$$
$$= [(POTW\_flow\_AD) \times (digester\ gas)/\ (per\ capita\ flow)] \times conversion\ to\ m^3 \times (FRAC\_CH_4) \times (365.25) \times (density\ of\ CH_4) \times (1\text{-}DE) \times 1/10^9$$

$$Total\ CH_4\ Emissions\ (kt) = A + B + C + D$$

primary treatment (67.5 percent), the maximum CH_4-producing capacity of domestic wastewater (0.6), and the relative CH_4 correction factor for well-managed aerobic (zero), not well managed aerobic (0.3), and anaerobic (0.8) systems with all aerobic systems assumed to be well-managed. The methodology is summarized below and full details can be found in Chapter 7-18, page 464. (https://www.epa.gov/sites/production/files/2017-04/documents/us-ghg-inventory-2016-main-text.pdf)

where,

US_{POP}	= U.S. population
% onsite	= Flow to septic systems / total flow
% collected	= Flow to POTWs / total flow
% aerobic	= Flow to aerobic systems / total flow to POTWs
% anaerobic	= Flow to anaerobic systems / total flow to POTWs
% aerobic w/out primary	= Percent of aerobic systems that do not employ primary treatment
% aerobic w/primary	= Percent of aerobic systems that employ primary treatment
% BOD removed in prim. treat.	= 32.5%
% operations not well managed	= Percent of aerobic systems that are not well managed and in which some anaerobic degradation occurs
% anaerobic w/out primary	= Percent of anaerobic systems that do not employ primary treatment
% anaerobic w/primary	= Percent of anaerobic systems that employ primary treatment
EF_{SEPTIC}	= Methane emission factor (10.7 g CH_4/capita/day) – septic systems
Days	= days per year (365.25)
Total BOD_5 produced	= kg BOD/capita/day × U.S. population × 365.25 days/yr
B_o	= Maximum CH_4-producing capacity for domestic wastewater (0.60 kg CH_4/kg BOD)
$1/10^6$	= Conversion factor, kg to kt
MCF-aerobic_not_well_man.	= CH_4 correction factor for aerobic systems that are not well managed (0.3)
MCF-anaerobic	= CH_4 correction factor for anaerobic systems (0.8)
DE	= CH_4 destruction efficiency from flaring or burning in engine (0.99 for enclosed flares)
POTW_flow_AD	= Wastewater influent flow to POTWs that have anaerobic digesters (MGD)
digester gas	= Cubic feet of digester gas produced per person per day (1.0 ft^3/person/day)
per capita flow	= Wastewater flow to POTW per person per day (100 gal/person/day)
conversion to m^3	= Conversion factor, ft^3 to m^3 (0.0283)
FRAC_CH_4	= Proportion CH_4 in biogas (0.65)
density of CH_4	= 662 (g CH_4/m^3 CH_4)
$1/10^9$	= Conversion factor, g to kt

The CH_4 and N_2O emissions of biological treatment can be estimated using the default method given in Equations 4.1 and 4.2 shown below:

EQUATION 4.1
CH_4 EMISSIONS FROM BIOLOGICAL TREATMENT

$$CH_4\ Emissions = \sum_i (M_i \bullet EF_i) \bullet 10^{-3} - R$$

Where:

CH_4 Emissions = total CH_4 emissions in inventory year, Gg CH_4

M_i = mass of organic waste treated by biological treatment type *i*, Gg

EF = emission factor for treatment *i*, g CH_4/kg waste treated

i = composting or anaerobic digestion

R = total amount of CH_4 recovered in inventory year, Gg CH_4

S4.2.5 Composting: Maasakkers et al. 2016 estimated emissions from composting sites using the IPCC Tier 1 methodology (IPCC 2006), which is the product of an emission factor and the mass of organic waste composted. Tier 1 equations for calculating CH_4 emissions from composting facilities are found here: 2006 IPCC Report Volume 5 Ch. 4, Pg. 4.5. (https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/5_Volume5/V5_4_Ch4_Bio_Treat.pdf)

S4.2.6 Petroleum: Maasakkers et al. 2016 estimated emissions from petroleum facilities using the 2014 EPA Greenhouse Gas Inventory (GHGI). Emissions are estimated for each activity by multiplying emission factors by the corresponding activity. Emission factors were gathered from the following sources: DrillingInfo (2015), “Methane Emissions from the Natural Gas Industry by the Gas Research Institute and EPA” (EPA/GRI 1996a-d), “Estimates of Methane Emissions from the U.S. Oil Industry” (EPA 1999), consensus of industry peer review panels, BOEMRE and BOEM reports (BOEMRE 2004, BOEM 2011), analysis of BOEMRE data (EPA 2005, BOEMRE 2004), and the GHGRP (2010 through 2014). Activity data were gathered from the following sources: DrillingInfo

(2015), the Energy Information Administration annual and monthly reports (EIA 1990 through 2015), (EIA 1995 through 2015a, 2015b), "Methane Emissions from the Natural Gas Industry by the Gas Research Institute and EPA" (EPA/GRI 1996a-d), "Estimates of Methane Emissions from the U.S. Oil Industry" (EPA 1999), consensus of industry peer review panels, BOEMRE and BOEM reports (BOEMRE 2004, BOEM 2011), analysis of BOEMRE data (EPA 2005, BOEMRE 2004), the Oil & Gas Journal (OGJ 2015), the Interstate Oil and Gas Compact Commission (IOGCC 2012), the United States Army Corps of Engineers, (1995 through 2015), and the GHGRP (2010 through 2014). Further details can be found in Chapter 3-61, Page 169. (https://www.epa.gov/sites/production/files/2017-04/documents/us-ghg-inventory-2016-main-text.pdf)

S4.2.7 Natural Gas: Maasakkers et al. 2016 estimated emissions from natural gas infrastructure using the 2014 EPA Greenhouse Gas Inventory (GHGI). This inventory breaks down the calculation into three steps. First, they calculate the potential CH_4 by collecting activity data on production and equipment in use and applying emission factors. Next, they calculate the amount of the CH_4 that is not emitted and use data on voluntary action and regulations. Finally, they calculate the net emissions by subtracting the CH_4 that is not emitted from the total CH_4 potential estimates to develop net CH_4 emissions. Further details can be found in Chapter 3-71, Page 179. (https://www.epa.gov/sites/production/files/2017-04/documents/us-ghg-inventory-2016-main-text.pdf).

Table S3.1 Kern County Vista-CA Dataset, AVIRIS-NG Source Data, Vista-CA Bottom-Up facility and sectoral CH_4 emissions breakdown. ** indicates sectors that were excluded due to no detection or attribution by AVIRIS-NG and GSAAM. All emissions are in Gg CH_4 a^{-1}.

IPCC Level 1	IPCC Level 3	Vista-CA Categories	Vista-CA Elements	AVIRIS-NG Source Data			Vista-CA Bottom-Up	
			Total # of Facilities	Total # of Facilities Flown	Total # of Facilities Estimated	Gg/year	Total # of Facilities Estimated	Gg/year
1. ENERGY	1A1 Energy Industries	Refineries	7	7	7	0.93	7	0.07
		Power Plants	43	33	33	2.27	35	0.29
		sub-total	***50***	***40***	***40***	***3.20***	***42***	***0.36***
	1B2 Oil & Natural Gas	Distribution Pipelines (segments)	12,092	8,822	2	1.18	0	0
		NG Fueling Stations	5	5	5	0.07	0	0
		NG Stations	142	93	93	10.08	142	12.86
		Oil & Gas Facilities	1,888	1,702	1702	87.63	1,888	0.76
		Oil & Gas Fields	101	68	68	14.16	101	43.91
		Oil & Gas Wells	142,534	132,820	**	0	142,534	0.00
		Pipelines (segments)	3,696	2,166	**	0	3,696	0.01
		Processing Plants	9	9	9	3.47	9	1.91
		Storage Fields	0	0	0	0	0	0
		sub-total	***160,467***	***145,685***	***1879***	***116.58***	***148,370***	***59.46***
		sub-total	***160,517***	***145,725***	***1919***	***119.79***	***148,412***	***59.81***
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 Enteric Fermentation	Feed Lots	2	2	2	0.14	2	0.09
	3A1 & 3A2 Enteric Fermentation & Manure Management	Dairies	53	35	35	6.43	44	59.03
	3A2 Manure Management	Digesters	5	4	4	0.22	5	0.85
		sub-total	***60***	***41***	***41***	***6.79***	***51***	***59.97***
4. WASTE	4A1 Managed Waste Disposal	Landfills	107	50	50	2.95	108	10.71
	4B Biological Treatment of Solid Waste	Composting Sites	18	6	6	0.16	18	0.76
	4D1 & 4D2 Domestic and Industirial Water Treatment & Discharge	Wastewater Treatment Plants	7	4	4	0.40	7	1.05
		sub-total	***132***	***60***	***60***	***3.51***	***133***	***12.52***
TOTAL			**160,709**	**145,826**	**2,020**	**130.09**	**148,596**	**132.29**

Table S3.2 Kern County Vista-CA dataset, CARB PMT, EPA FLIGHT, Maasakkers et al., and CALGEM facility and sectoral CH_4 emissions breakdown. All emissions are in Gg CH_4 a^{-1}.

IPCC Level 1	IPCC Level 3	Vista-CA Categories	Vista-CA Elements	CARB PMT		EPA FLIGHT		Maasakkers et al.	CALGEM
			Total # of Facilities	Total # of Facilities Estimated	Gg/year	Total # of Facilities Estimated	Gg/year	Gg/year	Gg/year
1. ENERGY	1A1 Energy Industries	Refineries	7	5	0.05	3	0.06		
		Power Plants	43	19	0.20	11	0.41		
		sub-total	*50*	*24*	*0.25*	*14*	*0.47*	*0.95*	*0.99*
	1B2 Oil & Natural Gas	Distribution Pipelines (segments)	12,092	0	0	0	0		
		NG Fueling Stations	5	0	0	0	0		
		NG Stations	142	1	0	1	0		
		Oil & Gas Facilities	1,888	18	3.91	5	0.14		
		Oil & Gas Fields	101	0	0	3	0		
		Oil & Gas Wells	142,534	0	0	0	0		
		Pipelines (segments)	3,696	0	0	0	0		
		Processing Plants	9	0	0	1	0		
		Storage Fields	0	0	0	0	0		
		sub-total	*160,467*	*19*	*3.91*	*10*	*0.14*	*191.15*	*31.41*
		sub-total	*160,517*	*43*	*4.16*	*24*	*0.61*	*192.10*	*32.40*
3. AGRICULTURE, FORESTRY & OTHER LAND USE	3A1 Enteric Fermentation	Feed Lots	2	0	0	0	0		
	3A1 & 3A2 Enteric Fermentation & Manure Management	Dairies	53	0	0	0	0		63.46
	3A2 Manure Management	Digesters	5	0	0	0	0		
		sub-total	*60*	*0*	*0*	*0*	*0*	*67.05*	*63.46*
4. WASTE	4A1 Managed Waste Disposal	Landfills	107	0	0	1	3.61	13.51	4.05
	4B Biological Treatment of Solid Waste	Composting Sites	18	0	0	0	0	0.76	
	4D1 & 4D2 Domestic and Industirial Water Treatment & Discharge	Wastewater Treatment Plants	7	0	0	0	0	1.08	1.71
		sub-total	*132*	*0*	*0*	*1*	*3.61*	*15.35*	*5.76*
TOTAL			**160,709**	**43**	**4.16**	**25**	**4.22**	**274.50**	**101.62**

VI. References

1. Lara. SB-1383 Short-lived climate pollutants: methane emissions: dairy and livestock: organic waste: landfills. 2015. California Legislative Information. https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB1383. (accessed 3rd May 2021).

2. Nunez and Pavley. AB-32 Air pollution: greenhouse gases: California Global Warming Solutions Act of 2006. California Legislative Information. 2006. https://leginfo.legislature .ca.gov/faces/billNavClient.xhtml?bill_id=200520060AB32. (accessed 5th May 2021).

3. Thurmond. AB-1496 Methane emissions (2015-2016). 2015. California Legislative Information. https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id= 201520160AB1496. (accessed 3rd May 2021).

4. Rafiq, T. Duren, R. M., Thorpe, A. K., Foster, K., Patarsuk, R., Miller, C. E., Hopkins, F. M. 2020. Attribution of methane point source emissions using airborne imaging spectroscopy and the Vista-California methane infrastructure dataset. Environ. Res., Lett. 15, 1-12, DOI: https://doi.org/10.1088/1748-9326/ab9af8.

5. Fischer, M. L. and Jeong, S. 2016. Evaluating the Bay Area Methane Emission Inventory. *Lawrence Berkeley National Laboratory*. LBNL Report #: LBNL-1006297. Retrieved from https://escholarship.org/uc/item/6db6d0gz.

6. Jeong, S., Newman, S., Zhang J., Andrews. A. E., Bianco, L., Bagley, J., Cui, X., Graven H., Kim J., Salameh, P., LaFranchi, B.W., Priest, C., Campos-Pineda, M., Novakovskaia, E., Sloop, C. D., Michelsen, H. A., Bambha, R. P., Weiss. R. F., Keeling, R., and Fischer, M. L. 2016. Estimating methane emissions in California's urban and rural regions using multitower observations, Journal of Geophysical Research: Atmosphere, 121, 031–13,049, DOI: doi:10.1002/2016JD025404.

7. Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R., Janssens-Maenhout, G. 2019. High resolution temporal profiles in the Emissions Database for Global Atmospheric Research (EDGAR). Nature Scientific Data, 7, DOI: https://doi.org/10.1038/s41597-020-0462-2.

8. Jeong, S., Hsu, Y.-K., Andrews, A. E., Bianco, L., Vaca, P., Wilczak, J. M., Fischer, M. L. 2013. A multitower measurement network estimate of California's methane emissions, J. Geophys. Res. Atmos., 118, 11339–11351, DOI: http://doi.org/10.1002/jgrd.50854.

9. Hsu Y.-K., VanCuren, T., Park, S., Jakober, C., Herner, J., FitzGibbon, M., Blake, D. R., Parrish, D. D. 2010. Methane emissions inventory verification in southern California. Atmospheric Environment., 44, 1 – 7, DOI: https://doi.org/10.1016/j.atmosenv.2009.10.002.

10. Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L. 2012. On the Sources of Methane to the Los Angeles Atmosphere. Environmental Science & Technology, 46, 9282-9289, DOI: https://doi.org/10.1021/es301138y.

11. Peischl, J., Ryerson, T. B., Brioude, J., Aikin, K.C., Andrews, A. E., Atlas, E., Blake, D., Daube, B. C., de Gouw, J. A., Dlugokencky, E., Frost, G. J., Gentner, D. R., Gilman, J. B., Goldstein, A. H., Harley, R. A., Holloway, J. S., Kofler, J., Kuster, W. C., Lang, P. M., Novelli, P. C., Santoni, G. W., Trainer, M., Wofsy, S. C., and Parrish, D. D. 2013. Quantifying sources of methane using light alkanes in the Los Angeles basin, California, J. Geophys. Res. Atmos., 118, 4974 – 4990, DOI: https://doi.org/10.1002/jgrd.50413.

12. Cui, Y. Y., Brioude, J., McKeen, S. A., Angevine, W. M., Kim, S.-W., Frost, G. J., Ahmadov, R., Pieschl, J., Bousserez, N., Liu, Z., Ryerson, T. B., Wofsy, S. C., Santoni, G. W., Kort, E. A., Fischer, M. L., and Trainer, M. 2015. Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin, J. Geophys. Res. Atmos., 120, 6698–6711, DOI: https://doi.org/10.1002/2014JD023002.

13. Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., and Blake, D. R. 2014. Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data, J. Geophys. Res. Atmos., 119, 7741–7756, DOI: https://doi.org/10.1002/2014JD021551.

14. Fairley D., and Fischer, M. L. 2015. Top-down methane emission estimates for the San Francisco Bay Area from 1990 to 2012. Atmos. Environ., 107, 9 – 15, DOI: http://doi.org/10.1016/j.atmosenv.2015.01.065.

15. California Air Resources Board. 2020. Pollution Mapping Tool Sources in Your Community. https://ww3.arb.ca.gov/ei/tools/pollution_map/pollution_map.htm. (accessed 13th, October 2021).

16. Jeong, S., Zhao, C., Andrews, A. E., Bianco, L., Wilczak, J. M., and Fischer, M. L. 2012. Seasonal variation of CH_4 emissions from central California, Journal of Geophysical Research: Atmosphere, 117, DOI: https://doi.org/10.1029/2011JD016896.

17. Maasakkers J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., Hight, C., DeFigueiredo, M., Desai, M., Schmeltz, R., Hockstad, L., Bloom, A. A., Bowman, K. W., Jeong S., and Fischer, M. L. 2016.Gridded National Inventory of U.S. Methane Emissions. Environmental Science & Technology. 50, 13123-13133, DOI: http://doi.org/10.1021/acs.est.6b02878.

18. Yadav, V., Duren, R., Mueller, K.,Verhulst, K. R., Nehrkorn, T., Kim, J., Weiss, R.F., Keeling, R., Sander, S., Fischer, M. L., Newman, S., Falk, Matthias, Kuwayma, T., Hopkins, F., Rafiq, T., Whetstone, J., and Miller, C. 2019. Spatio-temporally resolved methane fluxes from the Los Angeles megacity. Journal of Geophysical Research: Atmosphere, 124, DOI: https://doi.org/10.1029/2018JD030062.

19. Duren, R. M., Thorpe, A. K., Foster, K.T., Rafiq, T., Hopkins, F. M. Yadav, V., Bue, B. D., Thompson, D. R., Conley, S., Colombi, N.K., Frankenberg, C., McCubbin, I. B., Eastwood, M. L., Falk, M., Herner, J. D., Croes, B. E., Green, R. O., Miller, C. E. 2019. California's Methane 1 Super-Emitters. Nature 575, 180–184. https://doi.org/10.1038/s41586-019-1720-3.

20. Environmental Protection Agency. Facility Level Information on GreenHouse Gases Tool. 2018. https:// ghgdata.epa.gov/ghgp/main.do. (accessed 23rd May 2019).

21. Carranza, V., Rafiq, T., Frausto-Vicencio, I., Hopkins, F. M., Verhulst, K. R., Rao, P., Duren, R. M., and Miller, C. E. 2018. Vista-LA: Mapping methane-emitting infrastructure in the Los Angeles megacity. Earth Syst. Sci. Data, 10, 653-676, DOI: https://doi.org/10.5194/essd-10-653-2018.

22. Hopkins, F.M., Rafiq, T., and Duren, R.M. 2019. Sources of Methane Emissions (Vista-CA), State of California, USA. ORNL DAAC, Oak Ridge, Tennessee, USA, DOI: https://doi.org/10.3334/ORNLDAAC/1726.

23. Eggleston, H. S., L. Buendia, K. Miwa, T. Ngara, and K. Tanabe. 2006. IPCC Guidelines for National Greenhouse Gas Inventories. Prepared by the National Greenhouse Gas Inventories Programme, vol. 4. IGES, Japan.

24. Intergovernmental Panel on Climate Change. 2014. Climate Change 2014: Synthesis Report. https://archive.ipcc.ch/report/ar5/syr/. (accessed 3rd May 2021).

25. Thorpe, A. K., Roberts, D. A., Bradley, E. S., Funk, C. C., Dennison, P. E., Leifer, I. 2013. High resolution mapping of methane emissions from marine and terrestrial sources using a Cluster-Tuned Matched Filter technique and imaging spectrometry. Remote Sens. of Env., 134, 305 – 318, DOI: http://dx.doi.org/10.1016/j.rse.2013.03.018.

26. Thorpe, A. K., Frankenberg, C., Roberts, D. A. 2014. Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: application to AVIRIS. Atmos. Meas. Tech., 7, 491–506, DOI: https://doi.org/10.5194/amt-7-491-2014.

27. Thompson, D. R., Leifer, I., Bovensmann, H., Eastwood, M., Fladeland, M., Frankenberg, C., Gerilowski, K., Green, R. O., Kratwurst, S., Krings, T., Luna, B., and Thorpe, A. K. 2015. Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane. Atmos. Meas. Tech., 8, 4383–4397, DOI: https://doi.org/10.5194/amt-8-4383-2015.

28. Guha, A., Newman, S., Fairley, D., Dinh, T. M., Duca, L., Conely, S. C., Smith M. L., Thorpe, A. K., Duren, R. M., Cusworth, D. H., Foster, K. T., Fischer, M. L., Jeong, S., Yesiller, N., Hanson, J. L., and Martien, P. T. 2020. Assessment of Regional Methane Emission Inventories through Airborne Quantification in the San Francisco Bay Area. Environ. Sci Technol., 54, 9424 – 9264, DOI: https://doi.org/10.1021/acs.est.0c01212.

29. Hopkins, F. M., Kort, E. A., Bush, S. E., Ehleringer, J. R., Lai, C.-T., Blake, D. R., and Randerson, J. T. 2016. Spatial patterns and source attribution of urban methane in the Los Angeles Basin. J. Geophys. Res. Atmos., 121, 2490–2507, DOI: https://doi.org/10.1002/2015JD024429.

30. Krautwurst, S., Gerilowski, K., Krings, T., Borchard, J., Bovensmann H. 2016. Scientific and Technical Assistance for the Deployment of a flexible airborne spectrometer system during C-MAPExp and COMEX. DOI: https://doi.org/10.5270/esa-k85rkdq.

Chapter 4

Distribution of Methane Super-emitters Amongst California Power Plants

I. Introduction

Even with the progress in adopting alternative fuel technology and renewable energy resources, the energy industry is still a significant contributor to greenhouse gas (GHG) emissions. While most attention has been given to carbon dioxide (CO_2) emissions, energy industries also produce significant methane (CH_4) emissions, which are not as well understood. CH_4 is a relatively short-lived climate pollutant with respectively around 30 and 85 times the cumulative radiative forcing of CO_2 on a mass basis over 100-year and 20-year time frames.[1] Natural gas, which mostly consists of CH_4, is considered a "cleaner" fuel alternative, producing half the amount of CO_2 per unit of energy as coal, and has been considered a viable transitional resource until future renewable energy technologies mature.[2] While the transition from coal to natural gas has lowered overall GHG emissions from the electricity sector, it has also increased CH_4 emissions; over the last 30 years, CH_4 emissions in California from the electric generation sector have doubled.[3, 20] This increase in CH_4 emissions may be an underestimate as it does not completely account for fugitive emission occurring in the natural gas lifecycle. When fugitive CH_4 emissions are accounted for, the 20-year climate impact of natural gas increases by 50%, and can nullifying natural gas's benefits relative to other fuels.[4, 5] California's Assembly Bill 1496 and Senate Bill 1383 set CH_4 emissions reduction targets in order to reduce short-lived climate pollutants along with achieving its climate goals for significantly reducing greenhouse (GHG) emissions by 2030.[6, 7, 8] As such, the identification of CH_4 emission sources in electricity generation is important for effective development of appropriate policy and corresponding mitigation methods.

In California, close to 50% of all in-state electric generation comes from natural gas firing power plants (90.7 TWh).[3] CH_4 emissions at power plants are produced from the processes generating electricity (e.g., combustion of natural gas) and from fugitive leaks in the natural gas handling system. Process-based emissions result from the operations occurring at a given plant or are the byproduct of combustion activities that are usually persistent while the facility is operating. In contrast fugitive emissions are the result of leaks in the components or infrastructure of a given facility. To further classify different types of power plants, we organized power plants into two categories: stand-alone power plants (SPWP) and co-located power plants (CPWP). SPWP's house infrastructure dedicated to electric generation and on average they consume more fuel than CPWP's, leading to higher activity-based emissions estimation. We define CPWP's as power plants located or connected to an existing facility from different CH_4 emission sources that either provide power for operations for that facility or utilize captured CH_4 to generate power. For example, cogeneration power plants utilizing landfill gas are defined as being co-located with that specific landfill rather than being stand-alone power plants. CPWP's on average consume less fuel from other sources, given that they are designed to utilize local CH_4 sources such as landfills or dairy biogas.[2]

CH_4 inventories used by regulatory agencies quantify sectoral and facility emissions from power plants include process-driven emissions, such as from combustion of fuels, but lack accounting for fugitive emissions. However, fugitive emissions have been observed in many energy systems, particularly those that use natural gas as a fuel.[17, 18] Understanding the contribution of fugitive emissions is critical since they are more prevalent under direct observations and traditional methods of estimation do not account for them. Some approaches, such as IPCC tier 1 methods, use activity data, such as the amount of fuel burned, along with an emission factor to estimate power plant methane emissions. The Environmental Protection Agency (EPA) calculates CH_4 emissions from electric generation in the Greenhouse Gas Reporting Program (GHGRP) consistent with Intergovernmental Panel on Climate Change (IPCC) tier 1 methods.[33] IPCC tier 1 methods solely incorporate combustion processes by fuel type and do not include fugitive or venting emissions.[1]

Other regulatory inventories rely on data that is either reported by producers or measured at the stack. The U.S. EPA's Facility Level Information on GreenHouse gases Tool (EPA FLIGHT) tracks facilities that emit more than 25,000 $MTCO_2eq$/year, and accounts for 85-90% of emissions included in the official EPA GHGRP.[19] EPA FLIGHT differs from the official EPA GHG inventory because they are based on reported emissions at the facility scale, not activity data. EPA requires power plant facilities to install Continuous Emissions Monitoring Systems (CEMS) devices to measure CO_2, NO_x, and SO_x emissions from flaring or vent stacks but do not require the incorporation or the capability to measure emissions from fugitive leaks in the infrastructure.[19] Consequently, EPA FLIGHT's reported emissions do not include fugitive CH_4 emissions. Similarly, the California Air Resources Board's Pollution Mapping Tool (CARB PMT) records GHG and criteria pollutant emissions from large facilities in California that are subject to the GHG Mandatory Reporting Regulation (MRR).[20, 46] GHG MRR requires all electric generating facilities to self-report CH_4 emissions if they emit greater than or equal to 10,000

$MTCO_2eq$/year.[46] Vented and fugitive emissions are not required to be included in the GHG MRR.[46]

Uncertainty in both the quantification and source apportionment of California's CH_4 emissions hampers mitigation and convolutes emission reduction evaluations. There have been multiple studies focused on measuring CH_4 emissions from the production, storage, and processing within the energy and oil and gas industries, but little work has been conducted on understanding and quantifying fugitive CH_4 emissions from power plants in California. Some work looking at CH_4 emissions from power plants has been done in other parts of the U.S. Lavoie et al. studied 6 facilities, 3 natural gas-fired power plants and 3 refineries using top-down mass-balance approaches to quantify CH_4 and found resulting emissions to be higher than those reported to the EPA by factors of 21 – 120 for natural gas power plants.[17] Hajny et al. continued Lavoie et al.'s work by surveying 14 natural gas-fired power plants utilizing a similar mass-balance technique, compared CH_4:CO_2 ratios, and attributed emissions to un-combusted natural gas from power plant stacks rather than fugitive emissions, which again was unaccounted for in reported datasets.[18]

To understand these gaps in regulatory inventories, we first developed a bottom-up dataset of power plant CH_4 emissions using geospatial records of facility and infrastructure information for California coupled with fuel metrics and IPCC methodologies. Our bottom-up modeled dataset accounts for process emissions from fuel combustion, which produces CH_4 as a by-product of the incomplete combustion of fuel. We compared the modeled emissions to CH_4 plume observations made from airborne CH_4 imaging surveys of

California that have been extensively conducted over the last few years.[21 22, 23] We used CH_4 retrievals from the Next Generation Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-NG), which measures ground-reflected solar radiation in the 380 to 2,510 nm range with a field of view of 1.8 km and spectral sampling of 5 nm. At normal survey altitudes of 3 km AVIRIS-NG captures data at a 3 m pixel resolution and can consistently detect and measure CH_4 point sources with emissions typically as small as 2 – 10 kg CH_4 h^{-1} for typical surface winds $\leq$ 5 m s^{-1}.[24, 25, 26] AVIRIS-NG surveys in California were conducted from 2016 – 2018 and covered facilities in the waste, agriculture and energy sectors.[27] We also used CH_4 imaging from the Hyperspectral Thermal Emission Spectrometer (HyTES), a push-broom imaging spectrometer that produces a wide swath Thermal Infrared (TIR) image collecting spectral data from 256 bands ranging from 7.5 – 12 µm with a spatial resolution of around 2 m at a 1 km altitude.[28, 29, 30] Airborne surveys using the HyTES instrument targeting California oil and gas infrastructure were conducted from 2014 – 2017.[30] The meter-scale spatial resolution of both methane imagers enables attribution of methane plumes to source facilities (Rafiq et al., 2020), and also to individual infrastructure within a facility. By linking methane plume observations to infrastructure with a specific function, such as a chimney as the source of combustion emissions, we are able further attribute emissions to processes. Finally, we compare the modeled emissions with these airborne CH_4 plume observations to identify discrepancies in emission rates.

We hypothesize that fugitive emissions that do not emanate from power plant stacks are responsible for higher CH_4 emissions as compared to the publicly reported inventories. We further hypothesize that emissions we attribute to power plant processes are persistent while the power plant facility remains operational, and we expect to see these types of emissions exhibited periodically over the course of multiple top-down observations. We hypothesize that fugitive emissions to be one-off or emitted at a single time and due to their unpredictable nature, we also presume observing them once or twice given a facility is sampled multiple times. Moreover, we surmise higher overall CH_4 emissions from stand-alone power plants (SPWP) than co-located power plants (CPWP) as SPWP's house dedicated power generation infrastructure and consume more fuel on average than CPWP's. Overall, this study aims to show that fugitive emissions from power plants are underestimated in reported emission inventories which can affect potential mitigation strategies.

II. Methods

We used three bottom-up datasets to establish CH_4 emission estimates and identify locations for power plants in California: Vista-CA power plant geospatial dataset, 2017 CARB PMT facility data, and 2017 EPA FLIGHT facility data. Vista-CA is a geospatial database of 901,009 validated features of potential CH_4 emitting infrastructure in California for 2017 developed from publicly available datasets organized into the IPCC framework.[21, 22, 23] Vista-CA contains records for 433 power plants in California and categorizes them by their primary firing source: 114 are biomass, 2 are coal-powered, 283 are natural gas firing, 1 is nuclear-powered, 10 petroleum powered, and 23 are miscellaneous. Of those 433 power-plants, 329 are stand-alone power plants (SPWP) and 104 are co-located power plants (CPWP). In total, 104 power plants in Vista-CA are co-located with at least 1 or more of the following Vista-CA sectors: 42 power plants are co-located with landfills, 26 power plants co-located with wastewater treatment plants, 16 are co-located with oil and gas facilities, 1 is co-located with a dairy, 19 are co-located with refineries, and 3 are co-located with natural gas processing plants. CARB PMT data contained records for 592 facilities in California of which 227 were power plants (https://ww3.arb.ca.gov/ei/tools/pollution_map/).[20] EPA FLIGHT data contained records for 325 facilities in California of which 173 were power plants (https://ghgdata.epa.gov/ghgp/main.do).[19]

IPCC published guidelines for national GHG inventories in 2006 (https://www.ipcc-nggip.iges.or.jp/public/2006gl/index.html).[32, 33] These guidelines provide sector-specific methodologies for estimating greenhouse gas emissions using three tiers of increasing specificity for estimating CH_4 emissions for a given facility. We calculated power plant emissions using the guidance for stationary combustion (IPCC Level 1A Energy Industries), where stationary combustion refers to emissions from the intentional oxidation of materials within a power plant that is designed to raise heat and provide it either as heat or as mechanical work such as electricity. This study utilizes tier 2 stationary combustion method which solely requires multiplying fuel combustion statistics by country-specific emission factors derived from national fuel characteristics and do not include fugitive emissions.[32]

Equation 1.

$$E_f = EF_f \ x \ FC$$

Where E_f refers to the total CH_4 emissions by type of fuel (kg), EF_f refers to country-specific emission factor of a given GHG by type of fuel (kg gas/TJ), and FC refers to amount of fuel combusted (TJ).

The U.S. Energy Information Administration (U.S. EIA) annually publishes individual plant and generator data that includes metrics such as heat content of fuels and total fuel consumed. We obtained EIA plant information for 2019 which contained specific records of monthly and total fuel consumed for over 14,000 U.S. power plants. U.S.-specific fuel-based emission factors for 2018 were obtained from the Electronic Code of Federal Regulations (ECFR) via the EPA (https://www.ecfr.gov/cgi-bin/text-idx?SID=ae265d7d6f98ec86fcd8640b9793a3f6&mc=true&node=pt40.23.98&rgn=div5#ap40.23.98_138.2).[34] Next, we matched the fuel consumption power plant data from Vista-CA using the plant identification code, and we were able to match 321 power plants in total. Details on the emissions factors used can be found in the Supplementary Information (S1.1). Fuel type for each Vista-CA power plant was based on the primary firing source in the metadata. Finally, we used the IPCC Tier 2 equation to generate our bottom-up estimates in kg CH_4 a^{-1} for all 321 matched power plants.

Progress in spectroscopic instrumentation for top-down observations have allowed the detection and retrieval of CH_4 plumes at sub-meter scales across large areas. Here, we use CH_4 retrievals from two airborne infrared imagers, AVIRIS-NG and HyTES. The AVIRIS-NG instrument is unique in terms of its high signal to noise ratio, response uniformity, and calibration accuracy. Close to 2,000 flights conducted using the AVIRIS-NG instrument covered key CH_4 point source emission sectors including oil and natural gas production, processing, transmission, storage and distribution infrastructure. While AVIRIS-NG operates in the visible to shortwave spectrum and utilizes reflected solar radiance to detect CH_4,[24, 25, 26] HyTES' thermal infrared spectrometer relies on the thermal emission and thermal contrast between the ground and CH_4.[28, 29, 30] Close to 300 flights were conducted with the HyTES instrument over California, with multiple detections of CH_4 plumes. We used CH_4 plume observations data from AVIRIS-NG's and HyTES' California campaigns from 2014-2018.[23, 30] There were a total of 1,704 AVIRIS-NG plume observations with integrated CH_4 enhancement values (kg h^{-1}) collected during the 2016 – 2018 surveys.[23, 27] We linked quantified emission plumes to their source facilities and analyzed emission persistence to get a facility-level estimate of point source methane emissions across surveys for 795 unique methane sources of which included 24 power plant related sources.[16, 27] This method calculates the sum of all CH_4 plume emissions from a given facility and multiplies it by a persistence measurement defined as the ratio of observed plumes to the total number of AVIRIS-NG flights over a given area.[16] Additionally, the attribution algorithm, Geospatial Source Attribution Automated Model (GSAAM) was employed to geospatially associate AVIRIS-NG and HyTES plumes to power plants.[23] GSAAM attributes CH_4 point locations to facilities in the Vista-CA dataset through automated spatial association using a series of structured geospatial and logical relationships and sector-specific functions. Further details on the methods of this process can be found in Rafiq et al. 2020 (Supplementary S1.2).[23]

Subsequently, we obtained HyTES data and developed methods for vectorizing and operationalizing the hyperspectral imagery from each flight line for spatial analysis.[30] We collected the level 3 cluster-match-filtered CH_4 retrieval product (https://hytes.jpl.nasa.gov/order) for 252 flight-lines (July 2014 – June 2017). The available raw format of the hyperspectral data was in keyhole markup language zipped format (KMZ) which needed to be converted to a GeoTIFF format in order to execute operations. KMZ's were batch converted using QGIS' "translate" function. From there we developed a Python model to automate radiometric isolation of pixels highlighting CH_4 emissions for extraction (Supplementary S1.2). Next, all the extracted CH_4 emission GeoTIFF's were batch converted into polygon vectors using QGIS and then merged together to form one combined vector layer that identified 22,739 CH_4 emission polygons and associated metadata containing flight records for each plume. Finally, the centroid position of each polygon was taken in order to use this location within the GSAAM algorithm for HyTES plume attribution.[23] In order to derive HyTES flight coverage metrics, we used the 4th band in each HyTES multispectral GeoTIFF to extract and simplify flight outlines using a number of batch process functions and merged them all into one layer with associated flight metadata statistics (Supplementary S1.2).

Power plant attributions for AVIRIS-NG and HyTES were obtained using GSAAM. Power plant coverage for AVIRIS-NG was conducted using an assumption of 1800 m swath width across the centerline for each flight.[27] All Vista-CA power plant facilities lying inside this distance were considered to be included in the AVIRIS-NG survey. We repeated this process using the vectorized HyTES flight outlines to obtain coverage measurements. AVIRIS-NG plume points and HyTES plume polygons were directly overlaid on one another in order to spatially determine areas of overlap in California with a focus on power plant facilities. Finally, CH_4 emissions data from Vista-CA IPCC Tier 2 bottom-up emissions, AVIRIS-NG GSAAM source data, CARB PMT, and EPA FLIGHT were directly matched based on plant name and plant identification number for comparison and analysis. Finally, we explicitly identified CH_4 super-emitter influence in the power plant sector for all four datasets. Duren et al. defined sectors exhibiting CH_4 "super-emitter" influence where typically fewer than 20% of sources contribute to more than 60% of the total emissions from a given sector.[27] Vista-CA metadata and EIA metadata was used to establish the CH_4 emission relationship among SPWP's and CPWP's categories, fuel source type (biomass, petroleum, coal, natural gas, and other) and annual power generation (MWh) records.

We derived time series for all 37 power plants from which airborne methane plumes were observed to determine both the type of emissions possibly occurring and the degree to which they were persisting across flights or occurring as one-off fugitive emissions. All AVIRIS-NG CH_4 emissions attributed to 24 power plants through GSAAM were then further analyzed using geospatial methods and aerial imagery. We sought to identify sub-facility infrastructure to determine whether it was a fugitive leak or process emission and how often that piece of infrastructure was emitting, depending on the number of observations available for that given power plant facility. If a CH_4 plume was observed adjacent to a stack, we considered the emission to be related to fuel combustion at the plant (e.g., process emissions). If a CH_4 plume was observed elsewhere on the facility (e.g., near pipes, etc.), then we assumed it was a fugitive emission. Ambiguous cases occurred when plumes were found within the bounds of a power plant facility but were not tied to any particular piece of infrastructure or component. We utilized high-resolution 1 m aerial imagery from the U.S. Department of Agriculture's National Agriculture Imagery Program (NAIP) along with Google Earth's aerial imagery and street view images. Image analysis of the 24 AVIRIS-NG emitting power plants allowed us to pinpoint specific infrastructure in order to better determine emission type.

III. Results

The bottom-up and top-down survey of power plants presented here is the largest CH_4 emissions survey developed for California power plants. We calculated total expected CH_4 emissions of 2.76 Gg CH_4 a^{-1}, or 96.8% of the total in-state generation CH_4 emissions,[20] from the 321 out of 433 power plants in the Vista-CA dataset, accounting for 44.7 TWh of power (Figure 4.1 and Table 4.1).[3] Our power plant emission estimates revealed that natural gas power plants are expected to have a lower contribution to CH_4 emissions than biomass fueled plants, despite producing more power (Table 4.1). Of the 321 power plants modeled in Vista-CA, 0.7 Gg CH_4 a^{-1}, or 25% of total CH_4 was from 242 natural gas plants producing 88% of the power, while 2.04 Gg CH_4 a^{-1} was from 73 biomass fueled plants that produced close to 11% of the total power (Table 4.1).

CARB PMT and EPA FLIGHT data contain fewer power plants, with the Vista-CA power plant dataset cataloging 3 times and 6 times more power plants than the CARB PMT and EPA FLIGHT power plant datasets, respectively (Figure 4.1 and Table 4.2). The addition of more than 150 power plant facilities in the Vista-CA dataset over both CARB

PMT and EPA FLIGHT enables information on plants not tracked or surveyed before by these reporting agencies, which is critical for a thorough quantitative emissions assessment.

Vista-CA power plant CH_4 emissions exclude fugitive CH_4 given that IPCC Tier 2 methodology for power plant CH_4 only includes process emissions resulting from combustion of different fuels. Comparison of the Vista-CA emissions to CARB PMT and EPA FLIGHT shows a near 1:1 relationship, suggesting that similar methods were employed to calculate or report emissions, further suggesting that fugitive CH_4 emissions occurring at power plant facilities are not included in regulatory accounting of these facilities' CH_4 emissions (Figure S4.6A and S4.6B). As CARB PMT and EPA FLIGHT do not differ systematically from Vista-CA, it can be reasonably assumed that the modeled numbers are based on fuel consumption estimates.

AVIRIS-NG surveyed 253 power plants in California from 2016-2018 and observed 71 CH_4 plumes emitted from 24 power plants (Figure 4.2A, Tables 4.2 and S4.1A). Of the 24 emitting power plants, 8 power plants were SPWP's (all 8 natural gas firing) and 16 power plants were CPWP's (6 biomass, 5 natural gas firing, 5 other) (Tables 2 and S1A). Emission rates for individual plumes ranged from around 1 – 600 kg CH_4 h^{-1} (Table S1A). Comparatively, HyTES surveyed 30 power plants in California from 2014 – 2015 and observed 266 CH_4 plumes (via GSAAM) emitting from a total of 16 power plants (Figure 4.2B, Tables 4.2 and S4.1B). Of the 16 emitting power plants, 9 power plants were SPWP's (7 natural gas firing, 1 coal, and 1 other) and 7 were CPWP's (4 natural gas firing, 3 other) (Tables 2 and S1B). Only one GSAAM-attributed power plant was observed to be emitting by both AVIRIS-NG and HyTES (Figure S1A and S1B).

Using the CH_4 plume imagery, we then assessed the probable source of CH_4 emissions within each power plant, with a focus on determining whether emissions were from a fugitive or process-based origin. We analyzed imagery of CH_4 plumes from the 24 power plants observed to be emitting by AVIRIS-NG to determine the likely CH_4 emission source, which revealed 17 power plants emitting fugitive CH_4, and the remaining 7 plants producing process-based CH_4 (Table S1). CH_4 plumes we attributed to processes (Figures 4.4A, 4.4C, 4.4E) include plumes emanating from flaring heat stacks (Figures 4.4B, 4.4D, 4.4F), likely caused by the combustion process, and were observed from 7 different power plants. These observations of CH_4 from the same flaring heat stacks were repeatable in nearly half of facilities, with 4 observed emitting once and 3 observed emitting persistently in least 3 or more instances of observation of the same flaring heat stack. We also observed fugitive CH_4 plumes coming from pipelines (Figures 4.5A, 4.5B), storage tanks (Figures 4.5C, 4.5D), and natural gas related infrastructure (Figures 4.5E, 4.5F). Of the 17 AVIRIS-NG observed power plants showing fugitive leaks, 8 were one-time leaks, and 9 were showing signs of persistent activity with 5 of these 9 power plants seen emitting at least 3 more separate times from the same component.

Time series analysis of both HyTES and AVIRIS-NG data further confirmed that the observations attributed to processes were more persistent than those attributed to fugitive sources (Table S4.1). Process-attributed plumes were observed frequently over the course of multiple top-down observations while fugitive emissions were observed once or twice during multiple sampling periods. On average, fugitive plumes were observed 2-3 times more often than process-based plumes. We were able to derive more information using the time series for AVIRIS-NG since we had access to quantitative integrated CH_4 enhancements for detected plumes. AVIRIS-NG flights over the 24 GSAAM attributed power plants showed that the top 3 emitters, all of which were natural gas-firing, were only flown once and were observed to be emitting at a rate of at least 85 kg CH_4 h^{-1} (2 were classified as fugitive and one was process-based) (Table S1). 3 of the 24 AVIRIS-NG observed CPWP power plants were flown more than 12 times each and were observed emitting 8 or more CH_4 plumes (2 were classified as fugitive and one process-based, seen over multiple flights). It should be noted, there were temporal gaps spanning months between observation across most of those individual flights. Additionally, as stated in the methods, multiple plumes were observed in the same power plant facility from one flight were summed if those plumes were determined to be coming from the same origin or averaged if they were determined to be coming from different origins.

Further detailing the breakdown of fugitive and process- based emissions of the remaining 18 AVIRIS-NG-observed emitting power plants, 7 CPWP's (4 were fugitive and 3 were process-based) and 1 SPWP (process-based) were flown more than 10 times each and exhibited anywhere between 1 to 6 CH_4 plumes (Table S1A). 5 CPWP's (4 fugitive and 1 process-based) and 4 SPWP's (all fugitive) were flown at least 4 times and were observed producing 1-3 CH_4 plumes each (Table S1A). Only 1 biomass fueled CPWP was flown once and observed to be exhibiting fugitive emissions with a CH_4 plume observation of around 52 kg CH_4 h^{-1} (Table S1A). Table S1 outlines the above descriptions in more detail.

While emission rates are not quantified for HyTES, we were able to assess plume observation and attribution at the facility level. Of the 16 HyTES attributed power plants, 3 SPWP's (2 natural gas firing and 1 coal) and 2 CPWP's (2 natural gas firing) were flown more than 30 times with 181 attributed CH_4 plumes (Table S1B). 5 SPWP's (4 natural gas firing and 1 other) and 2 CPWP's (2 other) were flown more than 10 times with 70 attributed CH_4 plumes (Table S1B). Of the remaining 5 HyTES-attributed power plants that were flown more than 3 times with 15 attributed CH_4 plumes, 2 were SPWP's (3 natural gas firing) and 3 were CPWP's (2 natural gas firing and 1 other) (Table S1B).

With quantified CH_4 emissions observed by AVIRIS-NG for individual power plants, it is possible to compare observations to the expected emissions calculated using inventory methods. It should be noted that AVIRIS-NG instrument is unable to detect smaller emission sources below 2 – 10 kg h^{-1}, and is hence likely to underestimate overall emissions (Figures S4.4, S4.5).[27] Applying this sensitivity limit to the downscaled Vista-CA bottom-up emissions revealed 34 power plants expected to emit more than 2 kg CH_4 hr^{-1} and 9 plants emitting more than 10 kg CH_4 hr^{-1}. Simply put, scaling the AVIRIS-NG observations, we expect to see 1% or around 4 power plants to potentially exhibit emissions greater than 100 kg hr^{-1}, if all 433 Vista-CA power plants were surveyed in a day. Figure 4.3 illustrates this by showing the differences in cumulative percentages between power plants observed as emitting by AVIRIS-NG and the modeled rates from Vista-CA. AVIRIS-NG and Vista-CA vary from one another likely due to the occurrence of fugitive emissions at power plants that is unaccounted in the modeled estimates. Downscaling of Vista-CA emissions to the per hour mark (8,760 hours) greatly exemplifies the under accounting of emissions as compared to AVIRIS-NG (Figure S4.7). We further assumed a variety of down-scaling scenarios for CH_4 to be emitted a few hours out of the year (175 hours to 1,752 hours) to account for start-stop procedures and venting, and still found Vista-CA emissions to be lower than AVIRIS-NG emissions for the same facilities (Figure S4.7). On a per facility level, the differences in emissions between AVIRIS-NG and Vista-CA are more apparent. Comparing Vista-CA Bottom-Up power plants that overlap with AVIRIS-NG flight extents with AVIRIS-NG attributed power plants shows a difference in the distribution between biomass plants and natural gas firing plants (Table 1). For Vista-CA Bottom-Up, biomass plants emit 2.75 times more CH_4 than natural gas firing plants while AVIRIS-NG showed a reverse trend with natural gas firing plants emitting 2 times more CH_4 than biomass plants (Table 4.1). Both Vista-CA Bottom-Up and AVIRIS-NG modeled/observed at least twice as many natural gas firing plants as biomass plants (Tables 1, S1A, and S1B).

IV. Discussion

This work provides a first step towards assessing how remotely sensed observations of CH_4 plumes might be used to better understand actual CH_4 emissions from power plants. We hypothesized that there are undercounted fugitive emissions of CH_4 from power plants, and we showed that (a) regulatory datasets, such as EPA FLIGHT and CARB PMT, include fewer power plants than actually exist (b) power plant CH_4 emissions follow a heavy-tail distribution that suggests super-emitter behavior, and (c) attribution of individual CH_4 plumes to their likely sources supports the notion that fugitive emissions are undercounted, and may be responsible for a discrepancy between observations and modeled emissions.

Our newly calculated Vista-CA power plant CH_4 emissions dataset, along with remotely sensed CH_4 emissions from AVIRIS-NG provide the largest documented power plant CH_4 emissions dataset compiled for California (Table S4.3). These complimentary products illustrate a lack of relationship between power generation capacity and observed CH_4 emissions, despite being the core of currently used bottom-up models (Table 1). Both our bottom-up and top-down results show significant super-emitter behavior in the power plant sector (Figures S4.3, S4.4, and S4.5). We are able to show that AVIRIS-NG detected fugitive emissions emanating from power plants (Figure 4.5 and Table S4.1A). Both the bottom-up modeled emissions and the top-down measurements reinforced the notion of discrepancies existing in the assessment of fugitive emissions (Figure 4.3).

The total AVIRIS-NG emission estimate for all 24 facilities was 1,071 kg CH_4 hr^{-1}, of which half is derived from one natural gas stand-alone power plant facility (515 kg CH_4 hr^{-1}). In order to compare this figure with Vista-CA, we downscaled Vista-CA emissions, calculated annually, to hourly, assuming uniform emissions over the course of a year, resulting in a total of 311 kg CH_4 hr^{-1} for 321 power plants, which is 1.78-3.44 times smaller than AVIRIS-NG observations. Comparing across the same 24 plants between both datasets, AVIRIS-NG is at least 28-54 times larger than Vista-CA. Additionally, almost 90% of the 1,071 kg CH_4 hr^{-1} AVIRIS-NG observation was a result of fugitive-based emissions, while 137.4 kg CH_4 hr^{-1} was process-based. We upscaled every power plant emission observed by AVIRIS-NG, constrained it by potential operational hours of high emission and compared it with the total emissions estimate from Vista-CA. We saw that constraining AVIRIS-NG emissions to 10% high operation hours (876 hours resulting in

938,196 kg CH_4 a^{-1}) makes up about 35% of emissions from the total Vista-CA dataset (2,761,807.42 kg CH_4 a^{-1}). In other words, if power plants are assumed to be emitting 876 hours in a given year, AVIRIS-NG observations explain 35% of the total Vista-CA bottom-up inventory from only 24 AVIRIS-NG observed power plants. For context, this number increases to 70% explained emissions by AVIRIS-NG of total Vista-CA emissions with 20% assumed hours (1,752 hours) of operation time and almost 350% explained emissions by AVIRIS-NG of total Vista-CA emissions with 100% assumed hours (8,760 hours) of operation time. Similarly, Vista-CA downscaling under different hours of operation scenarios in a given year (175 hours to 1,752 hours) results in under accounting of AVIRIS-NG observed emissions by 1.5x to 10x. This suggests that emissions from the power plant sector seem to be underestimated in bottom-up models when compared to top-down observations.

On a per power plant facility level, natural gas firing power plants were observed to be emitting more CH_4 than biomass plants while the opposite was analyzed from the bottom-up product (Table 1). This result is also evident in the CARB GHG bottom-up inventory where biomass plants are modeled to out-emit natural gas plants 1.5:1.[20] This is likely due to the nature of the Tier I calculation for CH_4 emissions used, since burning natural gas has a lower emission factor than burning biomass. Because these models neglect fugitive CH_4, natural gas firing plants are not only producing more power on a MWh basis but are actually emitting more CH_4 on that basis as well. Vista-CA bottom-up revealed higher emissions for power plants that were detected to be less than 50 kg hr^{-1} as observed by AVIRIS-NG (Figure S4.5). Conversely, for the power plants detected by

AVIRIS-NG showing emissions greater than 50 kg hr^{-1}, the Vista-CA bottom-up product showed emissions that were negligible (Figure S4.5). Both of these notions can be related to temporal variability and not capturing peak emissions at the right time. Finally, both bottom-up and top-down datasets reconfirmed strong super-emitter behavior across the California power plant sector (Figures S4.3, S4.4, and S4.5).

Super-emitter behavior was evident for all power plant CH_4 emissions, but a heavier tailed distribution was exhibited in the AVIRIS-NG observations compared to the bottom-up and reported datasets (Figure S4.3 and S4.4). With 1.2% of the power plants surveyed responsible for 60% of total observed AVIRIS-NG power plant CH_4 emissions, this was the most skewed version of super-emitter influence out of all the datasets analyzed (Figure S3). Furthermore, looking at individual AVIRIS-NG power plant plume data, this super-emitter behavior was also observed, with 19 out of the 75 observed plumes constituting more than 60% of the observed AVIRIS-NG emissions (Figure S4.3). Vista-CA bottom up, CARB PMT, and EPA FLIGHT all exhibited super-emitter behavior to a lesser degree ranging from 5 – 18% contribution towards the 60% emissions cumulation mark (Figure S4.4). This suggests that CH_4 top-down observations can better constrain the super-emitter influence among these top emitting power plants better than the modeled versions. It also reaffirms that a handful of facilities emitting the greatest proportion of CH_4 can be more specifically targeted for CH_4 emissions reduction with a higher probability of success. There were two major spatial clusters for super-emitting power plants as identified by all four datasets in California: the western Kern County oil fields and the South Bay complex in Los Angeles. The relationship among the types of power plants, the fuel types consumed by these power plants, and the quantity of CH_4 emissions produced by the bottom-up and the top-down results illustrated that there are significant fugitive emissions of CH_4 from power plants, and that there are super emitters that exceed the expected emissions (Tables 4.1, 4.2, and S4.1A).

Time series and remote sensing analysis of AVIRIS-NG observations of power plants demonstrate that CPWP's are more likely to show infrequent smaller magnitude fugitive-based emissions while SPWP's are more likely to show significant or higher magnitude fugitive-based emissions (Table S1A). Simply put, SPWP's are spatially larger than CPWP's which means they potentially house more infrastructure. Consequently, the likelihood of any leakages or failures in these components increases simply with more of these components existing at a SPWP facility. In accordance with the bottom-up emission modeling, we observed that SPWP's produced higher CH_4 emissions than CPWP's. Vista-CA accounts SPWP's as producing 122.54MW of power per facility on average versus CPWP's 46.35MW. AVIRIS-NG observed SPWP's produced five times more CH_4 than CPWP's as observed by AVIRIS-NG, while Vista-CA Bottom-Up showed a doubling of CH_4 emission between SPWP's and CPWP's (Tables 1, S1A, and S3). Only 7 of the 24 observed power plants exhibited persistent process-driven emission patterns while 17 of the 24 observed exhibited fugitive-type emissions (Figures 4.2A and Table S4.1A). Additionally, only 2 of those 24 facilities showed infrequent fugitive-type emissions greater than 100 kg CH_4 hr^{-1}. This again speaks to the characteristics of persistence versus one-off fugitives with AVIRIS-NG revealing larger emissions only seen once at a given power plant while smaller or more process-based emissions are seen with more frequency

(Table S4.1A). Time series plots of one SPWP and one CPWP are provided in the Supplement that include both AVIRIS-NG observations and Vista-CA averages from the different emission scenarios discussed previously (Figure S4.7). Image analysis suggests fugitive leaks are much more common in power plants than process-based emissions and these fugitive leaks are more likely than not to be persistent or occurring with some frequency (Figure 4.5 and Table S4.1A). CPWP's were recorded as emitting more than SPWP's yet fugitive leaks dominate both types of plants. Emissions mitigation would be most effective in targeting and assessing fugitive leaks in CPWP's from storage tanks and natural gas related infrastructure and components.

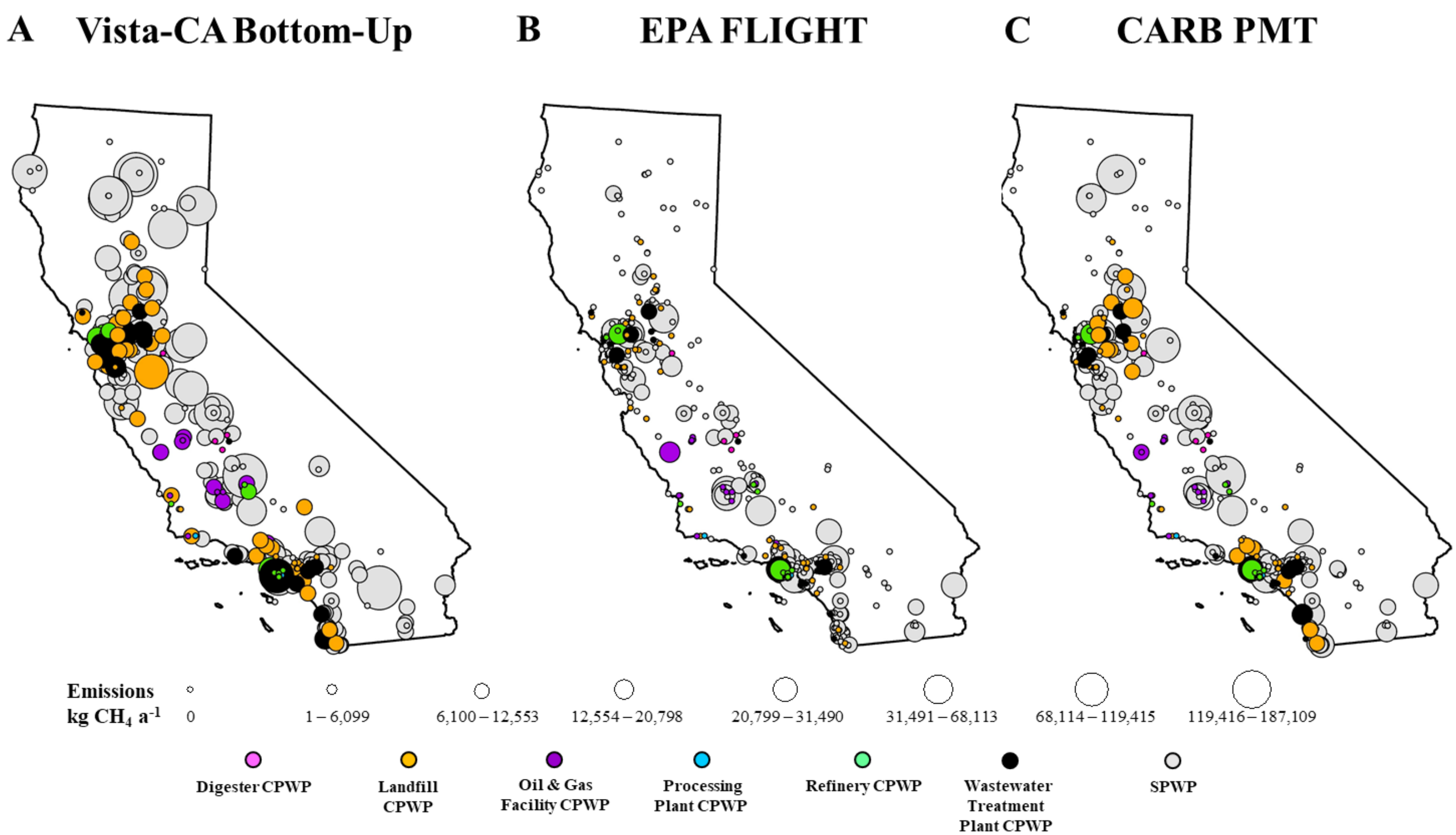
Limited number of observations by AVIRIS-NG and HyTES reduce the assessment efficacy for analyzing temporal variability of SPWP and CPWP CH_4 emissions and hinder high confidence determination of episodic and persistent categorization to CH_4 observations (Tables S4.1A and S4.1B). This also limits analyzing CH_4 emissions in the context of diurnal power generation variability which needs to be further investigated. While AVIRIS-NG and HyTES have significantly higher spatial resolution compared to existing satellites such as Tropospheric Monitoring Instrument (TROPOMI), Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), and Greenhouse Gases Observing Satellite (GOSAT), the temporal quality of aircraft observations can be further improved.[37, 38, 39] Dedicated flights over power plant facilities over an extended period of time can further reveal the frequency and probability of fugitive emission occurrence. Limitations on sustained observations hamper the potential to investigate fugitive emissions, which require a more committed platform to overcome these issues.[27] Facilities performing their own leak surveys on site periodically can be one way to address this. Additionally, ground-based surveys can also detect power plant CH_4 emissions, which are likely to be fugitive since a ground measurement would likely not be able to measure CH_4 coming from the stacks.

Multi-temporal focused flights using mass-balance methods or targeted/persistent aerial spectroscopy through high-resolution satellites is one way to overcome these data limitations and confidently answer critical questions about CH_4 emission dynamics. Current satellite instruments that provide capability for detecting GHG's such as CH_4 at higher resolutions include Greenhouse Gas Satellite (GHGSat), Precursore Iperspettrale della Missione Applicativa (PRISMA), and Environmental Monitoring and Analysis Program (EnMAP).[40, 41, 42, 43] Additionally, CH_4 dedicated high-resolution instrumentation is currently being developed in NASA's Geostationary Carbon Observatory (GeoCARB) and private ventures such as MethaneSAT.[44, 45] These instruments will potentially offer dedicated coverage of CH_4 producing infrastructure like power plants.

Modeled emissions using even IPCC Tier 2 specifics are still not complete as evidenced here by the discrepancy in relating emitter influence and fuel type from both the bottom-up and top-down perspectives. To improve bottom-up estimates, Tier 3 methods can be employed by incorporating specifics on combustion technology, control technology, quality of maintenance and the age of equipment utilized at each power plants.[32] Data at this level of detail is not only publicly inaccessible, but the quality of data is completely dependent on a per facility basis. Not every facility contains timely records or accurate records and data restrictions further complicated very detailed emissions assessment. IPCC also recommends utilizing Tier 3 methodologies but also acknowledges the difficulty in collecting very detailed facility data.[32] Power plants should have facility CH_4 leak surveys as part of their emissions reporting protocol as no IPCC method will be able to account for fugitive CH_4 emissions. Those measurements can be incorporated into new improved regulatory products that keep track of actual emissions (not just calculated process emissions).

We analyzed California's power plants through the context of top-down observations and bottom-up modeling techniques. Differences among datasets can yield uncertainty in designing actionable applications for CH_4 mitigation. Future work should implement persistent measurements to completely capture power plant CH_4 fugitive emissions throughout their operational cycle. As shown here, a complementary approach can be ideal for accurate constraining of emissions to benefit policy informing, especially as they relate to power plants.

Figure 4.1 Power plant locations in California, where the emissions are represented by the size of the marker, and type of power plant by the color, from datasets A) Vista-CA Bottom Up, B) EPA FLIGHT, and C) CARB PMT.



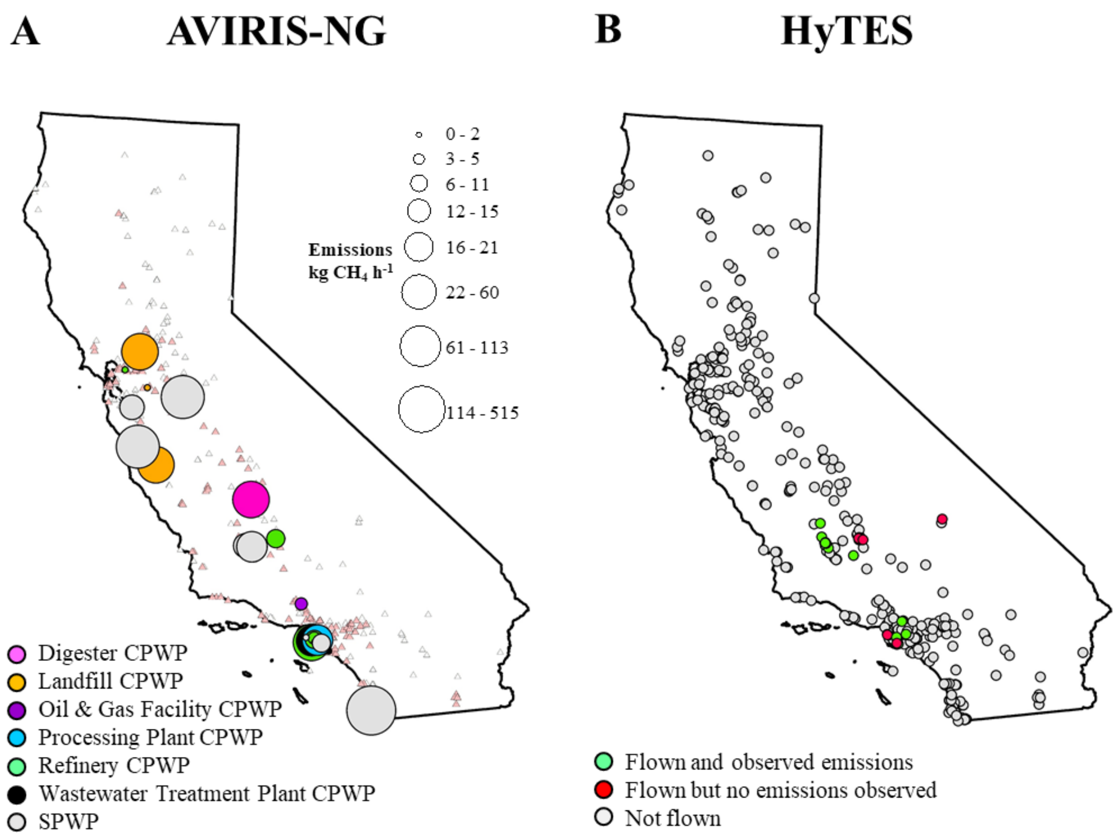


Figure 4.2 A) Power plant CH_4 emissions derived from AVIRIS-NG surveys (kg CH_4 h^{-1}). Circles represent plants with observed CH_4 emissions, with the size indicating emission magnitude, pink triangles indicate flown power plants with no observed emissions and grey triangles indicate power plants not flown by AVIRIS-NG) and B.) Power plant CH_4 observations from HyTES surveys.

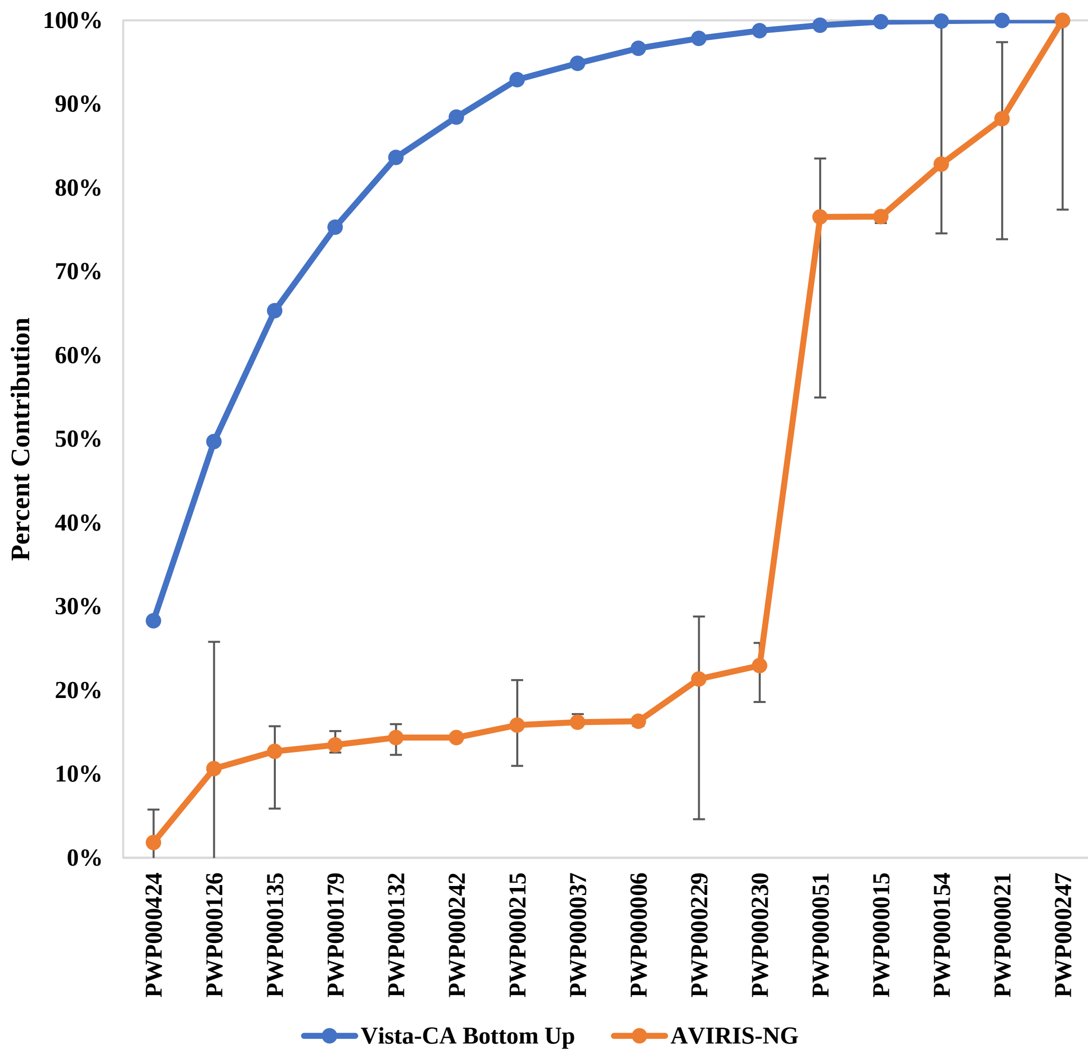


Figure 4.3 Vista-CA (blue) and AVIRIS-NG (orange) power plant cumulative distribution comparison for each of the 16 matched power plant facilities from each dataset. For all the power plants where CH_4 plumes were quantified by AVIRIS-NG, we totaled up the contribution of all power plants and divided the emissions from each power plant by that total to get the fractional contribution of each facility to the observed emissions as a percentage (orange bar). We chose the identical power plants from the Vista-CA Bottom-Up dataset and performed the same procedure (blue bar) and calculated the uncertainty for AVIRIS-NG using Duren et al.'s data.

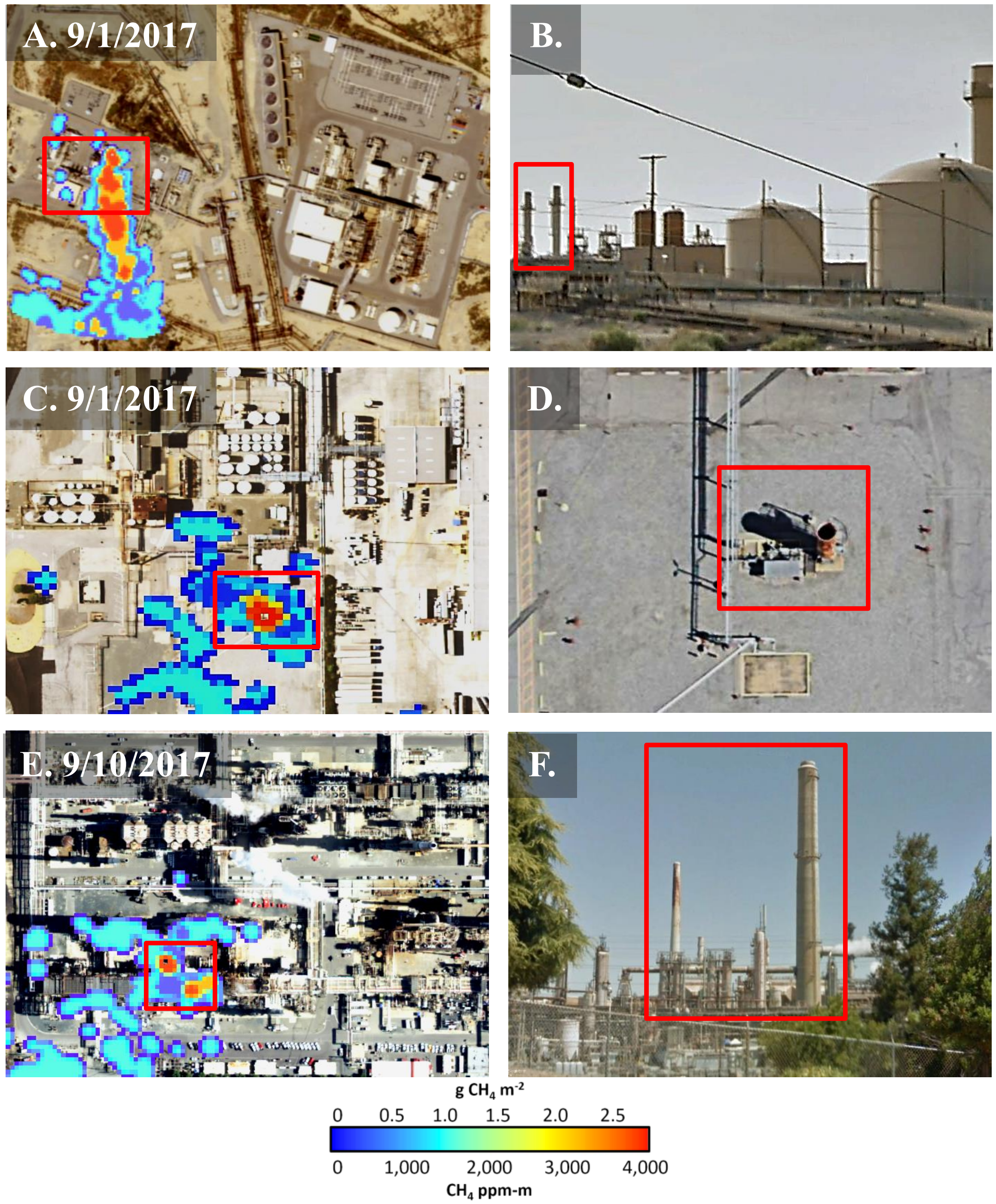


Figure 4.4 Power plant AVIRIS-NG CH_4 survey images illustrating process-based emissions emanating from flaring stacks. Images on the left show the AVIRIS-NG plume image overlaid on NAIP high-resolution imagery. Images on the right show the Google Earth aerial imagery/street-view. The color scales show the CH_4 concentration enhancement (the mass of CH_4 in a plume relative to background air) in each pixel.

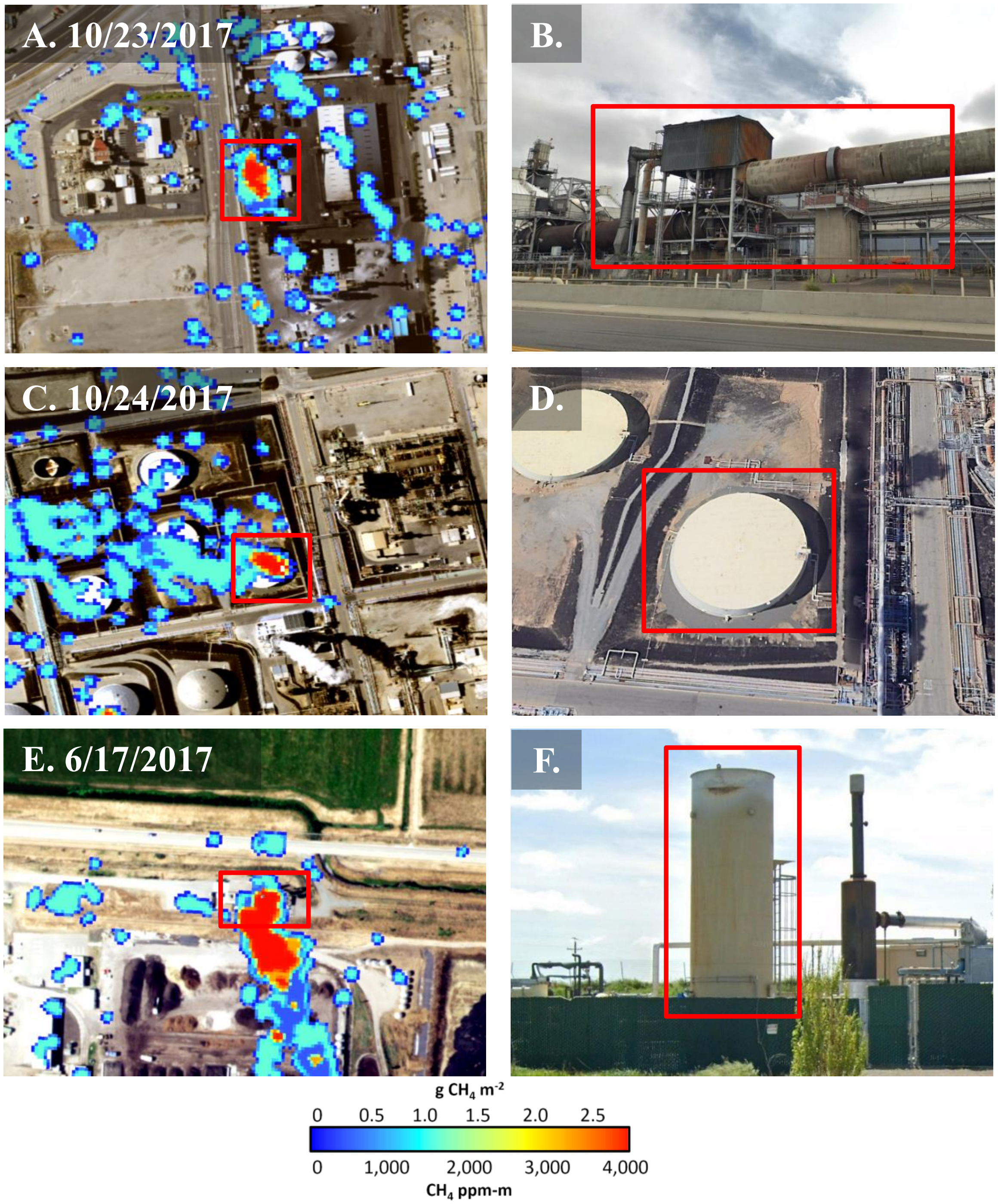


Figure 4.5 Power plant AVIRIS-NG CH_4 survey images showing fugitive emissions emanating from leaks in pipelines, storage tanks, and possibly ancillary components located in power plants. Images on the left show the AVIRIS-NG plume image overlaid on NAIP high-resolution imagery. Images on the right show the Google Earth aerial imagery/street-view. The color scales show the CH_4 concentration enhancement (the mass of CH_4 in a plume relative to background air) in each pixel.

Table 4.1 Breakdown of Vista-CA modeled emissions based on fuel source and power generation amount and AVIRIS-NG observations based on GSAAM attribution of CH_4 plumes to power plants. Third column identifies all Vista-CA Bottom-Up power plants where AVIRIS-NG flew.

Fuel Source	Vista-CA Bottom Up 321 Power Plants (kg CH_4 a^{-1})	Vista-CA Bottom Up (AVIRIS Survey Extent) 184 Power Plants (kg CH_4 a^{-1})	Vista-CA Bottom Up (AVIRIS Survey Extent) Per Power Plant (kg CH_4 a^{-1} per plant)	AVIRIS-NG Survey 24 Power Plants (kg CH_4 hr^{-1})	AVIRIS-NG Survey Per Power Plant (kg CH_4 hr^{-1})	EIA Power Generation (MWh)
Biomass	2,039,297.41 (73 plants)	452,514.32 (41 plants)	11,036.93	190.75 (6 plants)	31.79	4,668,389.27
Coal	12,552.60 (1 plant)	-	-	-	-	236,501.21
Natural Gas	703,131.16 (242 plants)	572,044.45 (140 plants)	4,086.03	796.77 (13 plants)	61.29	39,226,556.83
Petroleum	457.43 (1 plant)	-	-	-	-	12.00
Other	6,368.82 (4 plants)	6,283.07 (3 plants)	2,094.36	84.78 (5 plants)	16.96	544,770.39
TOTAL	**2,761,807.42**	**1,030,841.84**		**1,072.30**		**44,676,229.71**

Table 4.2 Number of Vista-CA power plants by primary fuel source surveyed by AVIRIS-NG and HyTES using both geospatial coverage and attribution of CH_4 plume observations to power plants using GSAAM attribution along with number of surveyed plants found in CARB PMT and EPA FLIGHT datasets

Primary Fuel Source	**AVIRIS-NG Survey Coverage**	**AVIRIS-NG Plants with CH_4 Observed**	**HyTES Survey Coverage**	**HyTES Plants with CH_4 Observed**	**CARB PMT Matched Vista-CA Power Plants**	**EPA FLIGHT Matched Vista-CA Power Plants**
Biomass	66	6	1	0	26	0
Coal	0	0	1	1	0	0
Natural Gas	170	13	24	11	96	77
Other	14	5	4	4	0	0
Petroleum	3	0	0	0	0	0
TOTAL	**253**	**24**	**30**	**16**	**122**	**77**

V. Supplementary

S4.1. Introduction

This document outlines the specific steps and procedures taken to develop and construct the data, methods, and analysis summarized in the main paper. The Vista-CA dataset is available at https://doi.org/10.3334/ORNLDAAC/1726. Attribution products from the Geospatial Source Attribution Automated Model (GSAAM) are available by request (https://doi.org/10.1088/1748-9326/ab9af8). The AVIRIS-NG dataset is available at https://data.carbonmapper.org. California Air Resources Board Pollution Mapping Tool (CARB PMT) dataset and information can be found at https://ww3.arb.ca.gov/ei/tools/pollution_map/. U.S. Environmental Protection Agency's Facility Level GreenHouse gas Tool (EPA FLIGHT) dataset and information can be found at https://ghgdata.epa.gov/ghgp/main.do#.

S4.1.1 Bottom-Up Data Collection and Generation

- CARB PMT geospatial data downloaded from https://ww3.arb.ca.gov/ei/tools/pollution_map/pollution_map.htm
    - Extracted 122 power plant records from this dataset
- EPA FLIGHT geospatial data downloaded from https:// ghgdata.epa.gov/ghgp/main.do
    - Extracted 78 power plant records from this dataset
- Vista-CA power plants dataset (Hopkins et al. 2019, Rafiq et al. 2020)
    - Use IPCC 2006 Stationary Combustion Tier 2 methodology for calculation
    - Match fuel consumption emissions data with vista-ca power plant data
    - Matched 321 power plants total
- Bottom-up emissions generation
    - Fuel consumption values from EIA multiplied by custom emission factors for IPCC 2006 Stationary Combustion Tier 2 Emissions
        1. Found fuel consumption data from EIA (2018)

a. Matched fuel consumption data from EIA to existing Vista-CA power plant dataset
b. Of the 433 power plants in Vista-CA, able to match 321 power plants with EIA fuel consumption data
c. Converted Fuel consumption units from MMBTU to TJ

2. Use default emission factors from IPCC tier 2 Stationary Combustion information, using the "PrimSource" field in Vista-CA to target and assign values (e.g. Biomass, Wood, Municipal Waste, Petroleum Coke, Petroleum Liquid, Coal, Natural Gas) in units of Kg CH_4/TJ
3. Multiplied fuel consumption data with emission factor to generate unique power plant emissions

S4.1.2 Airborne Dataset Collection and Processing

NASA JPL Next Generation Airborne Visible Infrared Imaging Spectrometer (AVIRIS-NG):

- AVIRIS-NG/Vista-CA GSAAM (Rafiq et al. 2020 and Rafiq et al. 2021 *in prep*)
    - Used attribution data to identify 8 power plants from 71 detected CH_4 plumes
    - Utilized source data development methods from previous paper to get facility emissions
        - The direct observational data method utilizes the sum of all fugitive and hotspot plume emissions from a given facility and multiplies it by a persistence measurement which is the ratio of the observed plumes to the total number of overflights. Normalizing by the persistence allows for us to potentially account for various temporal characteristics in facility observations.
        - **Equation 1.**

$$\boldsymbol{Q_{TU}} = \sum(\boldsymbol{\bar{Q}_n} \times \boldsymbol{P_n}), \qquad \boldsymbol{where\ P} = \left(\frac{\boldsymbol{O}}{\boldsymbol{T}}\right)$$

        - Where $\boldsymbol{Q_{TU}}$ is the total facility emissions estimate, $\boldsymbol{Q}$ is the individual plume estimated by AVIRIS-NG (Duren et al. 2019), $\boldsymbol{P}$ is the persistence, $\boldsymbol{n}$ is the source number, $\boldsymbol{O}$ is the total number of observed plumes, and $\boldsymbol{T}$ is the total number of overflights for a given facility.

NASA JPL's Hyperspectral Thermal Emissions Spectrometer:

- HyTES Emissions Extraction/Isolation
    - Collected data for 295**(252 operational) flight-lines starting from July 2014 – June 2017 (43 flight lines contained empty/broken data)
    - Downloaded 295 flight line data as KMZ's from JPL HyTES website: https://hytes.jpl.nasa.gov/order
        1. Kmz to tiff batch conversion using "Translate" function in QGIS
        2. Developed custom remote sensing model to automate radiometric isolation of emissions from TIFF into polygons using ArcGIS Model Builder/Python/SQL
            a. Parse tiffs into individual bands
            b. Convert individual bands to GeoTIFF's
            c. Re-classify individual band tiffs accordingly
                i. Band 1 reflectance values < 100
                    1. SQL code: {0-99} = 1, {100-255, No Data} = 2
                ii. Band 2 reflectance values > 100
                    1. SQL code: {0-99, No Data} = 2, {100-255} = 1
                iii. Band 3 reflectance values < 100
                    1. SQL code: {0-99} = 1, {100-255, No Data} = 2
            d. Sum Bands 1-3
                i. SQL code: {Band 1 + Band 2 + Band 3}
            e. Extract cells
                i. SQL code: {value = 3}
        3. Batch process Tiff to polygon conversion using the "Polygonise" function in QGIS
        4. Merge all polygons in QGIS using "Merge Vector Layer" function
        5. Final vector polygon data contains footprints for 22,739 emissions
        6. Calculate centroid of polygons and converted into points using "polygon to points" function
        7. Run through GSAAM for source attribution
        8. Extract HyTES footprints (batch automated processing):
            a. GDAL Batch polygonise function using the Alpha Band (Band 4)
            b. QGIS Batch Extract by attributes DN > 0
            c. GDAL Batch Dissolve function

d. QGIS Batch Delete Holes function
e. QGIS Batch Merge Function

S4.2. Analysis Methods

S4.2.1 Spatial Analysis

Spatial analysis between AVIRIS, HyTES, Vista-CA, EPA FLIGHT, CARB PMT for power plants using geospatial methods: spatial overlap, clip, intersect, and select by location

- Vista-CA IPCC Tier 2 compared with EPA FLIGHT data
    - Matched 77 Vista-CA power plants to EPA FLIGHT
        - All are Natural Gas
- Vista-CA IPCC Tier 2 compared with CARB PMT data
    - Matched 122 Vista-CA power plants to CARB PMT
        - 26 Biomass, 96 Natural Gas
- 253 Total Power Plants surveyed by AVIRIS (66 Biomass, 170, Natural Gas, 14 are other, 3 petroleum)
    - 8 Power plants attributed to AVIRIS plumes using GSAAM
        - All 8 are natural gas power plants
- At least 30 Total power Plants surveyed by HyTES
    - At least 9 Vista-CA power plants attributed to HyTES plumes using GSAAM
        - 7 are natural gas power plants, 1 is coal power plant, 1 is other type power plant
    - At least 7 Vista-CA Power Plants sub-facility attribution to HyTES using GSAAM
        - (All are originally attributed to refineries) 4 are natural gas power plants, 3 are other type power plant
    - At least 16 Vista-CA power plants attributed to HyTES plumes (100m buffer)
        - Power plants in Bakersfield, Long beach and El Segundo
- At least 188 AVIRIS point plumes spatially match with HyTES polygon plumes (100m buffer)
    - 5 “Sub-facility” power plants
- At least 552 HyTES polygon plumes spatially match with AVIRIS point plumes (100m buffer)

- At least 571 HyTES polygon plumes spatially match with Vista-CA Power Plants (100m buffer)
- No direct match in power plant data between AVIRIS attributed power plants and HyTES attributed power plants (El Segundo exception through sub-facility power plant in refinery)
- 64 Power plants have Vista-CA IPCC Tier 2 data, EPA FLIGHT data, CARB PMT data
    - 45/64 were flown by AVIRIS-NG (All are natural gas)
        - 5/64 observed to be emitting by AVIRIS-NG (All are natural Gas)
    - At least 4/64 were flown by HyTES (All are Natural gas)
        - 2/4 were seen to be emitting by HyTES (All are Natural gas)

S4.2.2 Statistical Analysis

All histograms show what we have seen with AVIRIS-NG that only a select number of power plants are strong emitters; the rest are smaller or even negligible. There is a clear systemic relationship happening when looking at the CO_2/CH_4 ratios for EPA FLIGHT and CARB PMT.

(SPWP vs CPWP: Big enough leak to detect by AVIRIS-NG, CPWP smaller than SPWP via histogram)

S4.3. Power Plant Results

S4.3.1 Power Plant Emitters and Fuel Consumption Type

- Vista-CA Bottom Up Emissions (316 facilities): Total 2,761,807.42 kg CH_4 a^{-1}
    - Biomass: 2,039,297.41 kg CH_4 a^{-1} (73.84%) (10.45% power generated)
    - Natural Gas: 703,131.16 kg CH_4 a^{-1} (25.46%) (87.80% power generated)
    - Petroleum: 457.43 kg CH_4 a^{-1} (0.02%) (0.000027% power generated)
    - Other: 6,368.82 kg CH_4 a^{-1} (0.23%) (1.22% power generated)
    - Coal: 12,552.60 kg CH_4 a^{-1} (0.45%) (0.53% power generated)
- AVIRIS-NG Source Data (normalized by amount of times flown)
    - 24 SPWP and CPWP: 1073.31 kg of CH_4 hr^{-1}
        - 13 Natural Gas: 796.77 kg of CH_4 hr^{-1} (74.23%) (87.80% power generated)
        - 6 Biomass: 190.75 kg of CH_4 hr^{-1} (17.78%) (10.45% power generated)

        - 5 Other: 84.78 kg of CH_4 hr^{-1} (7.89%) (1.22% power generated)

S4.3.2 Net Electricity Generation/Vista-CA Survey Context

- Total Vista-CA Bottom-Up Plants: 321
    - We have NETGEN info for 316 plants (The other 5 were Natural Gas Plants also)
- Total Power generated in 2019 from the 316 plants: 44,676,229.71 MWh
    - 237 Natural Gas Plants: 39,226,566.83 MWh (87.8%)
    - 73 Biomass Plants: 4,668,389.274 MWh (10.45%)
    - 1 Petroleum: 12 MWh (0.000027%)
        - (Children's Hospital) Natural Gas and Distillate Fuel Oil. Including diesel, No. 1, No. 2, and No. 4 fuel oils.
    - 4 Other: 544,770.39 MWh (1.22%)
    - 1 Coal: 236,501.21 MWh (0.53%)
- Comparison to Total In-State System Generation: 194,842,000 MWh (CEC 2018)
    - Natural Gas: 90,691,000 MWh (46.5%)
    - Biomass: 5,909,000 MWh (3.03%)
    - Petroleum: 430,000 MWh (0.22%)
    - Coal: 294,000 MWh (0.15%)

S4.3.3 HyTES Power Plant Detection Breakdown

- HyTES observations have come from plants that are completely natural gas-firing or partially natural gas fueled
- SPWP:
    - 1 Coal: PWP000029 Argus Cogen Plant (Conventional Steam Coal), concluded operations in 2014 (CEC), HyTES flew and saw emissions observed on July 8, 2014
    - 1 Other: PWP000046 BP Carson Refinery (Natural Gas and Other Gases)
        - Is also a co-located power plant as well
- CPWP:
    - HyTES surveyed 13 total power plants (All 13 have all or some Natural Gas usage)
        - 6 of 13 were not observed emitting by HyTES (using GSAAM):
            - 5 Natural Gas
            - 1 Other:
                - PWP000458 Wilmington Hydrogen Plant (natural gas, Waste Heat not directly attributed to a fuel source)
        - 7 of 13 HyTES observed emitting facilities (using GSAAM) (All co-located with refineries)
            - 4 Natural Gas:

- PWP000121 Dominguez Plant, (Natural Gas), co-located with SHELL OIL PRODUCTS
- PWP000132 El Segundo Cogen, (Natural Gas), co-located with CHEVRON USA INC
- PWP000178 Harbor Cogen, (Natural Gas), co-located with VALERO REFINING CO CALIFORNIA
- PWP000285 Oildale Energy LLC, (Natural Gas), co-located with Tricor Refining LLC

- 3 Other:
  - PWP000046 BP Carson Refinery (Natural Gas and Other Gases), co-located within TESORO REFINING & MARKETING CO
  - PWP000141 Equilon Los Angeles Refining (Other Gas), Secondary fuels: natural gas, Type: Burner, turbine
  - PWP000420 Tesoro Wilmington Calciner (Other primary fuel), Natural Gas, Petroleum Coke, Waste Heat (not attributed to fuel source)

S4.3.4 AVIRIS-NG Power Plant Detection Breakdown

AVIRIS-NG SPWP:

- Total flown: 168 stand-alone power plants
  - 160 non-emitting as determined by GSAAM
  - 8 emitting as determined by GSAAM
    - All 8 are natural gas plants

AVIRIS-NG CPWP:

- Total flown: 85 co-located power plants
  - 68 not observed to be emitting by AVIRIS-NG
  - 17 were observed to be emitting by AVIRIS-NG (9/17 have all or some Natural Gas usage)
    - 6 Biomass
      - 3 co-located with Landfills
        - PWP000015 Altamont Gas Recovery, Biomass, co-located with Altamont Landfill & Resource Recovery, (Landfill Gas)
        - PWP000021 Ameresco Johnson Canyon, Biomass, co-located with Johnson Canyon Sanitary Landfill, (Landfill Gas)
        - PWP000154 G2 Energy Hay Rd, (Biomass) co-located with Recology Hay Road Landfill, (Landfill Gas)
      - 1 co-located with Dairy

        - PWP000305 Pacific Rim Dairy Digester, (Biomass, Other Waste Biomass), co-located with Pacific Rim Dairy
    - 2 co-located with WWTP
        - PWP000323 Plant No 2 Orange County, (Biomass), co-located with ORANGE COUNTY S.D. #2 WWTP
        - PWP000424 Total Energy Facilities, (Biomass, Other Waste Biomass, Other Biomass Gas. Including digester gas, methane, and other biomass gases.), co-located with LACSD Joint WPCP WWTP
- 6 Natural Gas
    - 4 co-located with Refineries
        - PWP000121 Dominguez Plant, (Natural Gas), co-located with SHELL OIL PRODUCTS
        - PWP000132 El Segundo Cogen, (Natural Gas), co-located with CHEVRON USA INC
        - PWP000242 Martinez Refining, (Natural Gas, Oil and Gas), co-located with Shell Oil Products US – Martinez
        - PWP000285 Oildale Energy LLC, (Natural Gas), co-located with Tricor Refining LLC
    - 2 co-located with Oil and Gas Facilities
        - PWP000037 Berry Placerita Cogen, (Natural Gas), co-located with Central Facilities Oil and Gas Facility, (Combo Gas/Oil)
        - PWP000135 Elk Hills Power LLC, (Natural Gas), co-located with Occidental of Elk Hills, Inc.
- 5 Other
    - 4 co-located with Refineries
        - PWP000046 BP Carson Refinery, (Natural Gas and Other Gases), co-located with TESORO REFINING & MARKETING CO
        - PWP000229 Los Angeles Refinery Wilmington, (Natural Gas, Oil and Gas), co-located with Philips 66 COMPANY
        - PWP000420 Tesoro Wilmington Calciner, (Other, Natural Gas, Petroleum Coke, Waste Heat (not attributed to fuel source), co-located with VALERO REFINING CO CALIFORNIA
        - PWP000145 ExxonMobil Oil Torrance Refinery, (Other), co-located with EXXONMOBIL REFINING & SUPPLY CO
    - 1 co-located with NG Processing Plant
        - PWP000388 Signal Hill West Unit, (Other, Other Gases), co-located with West Unit Oil and Gas Facility

and SIGNAL HILL WEST UNIT GAS PLANT NG Processing Plant

S4.3.5 Super Emitter Characteristics

Roughly 10% of sources produce 60% of emissions; "Where typically fewer than 20% of sources (so-called super-emitters) contribute more than 60% of total emissions from that sector (Duren et al. 2019)", As we can see, power plants, both observationally and based on modeling, follow the super-emitter concept to a higher degree.

- AVIRIS-NG Source Data (253)
    - 1.2% (3) of power plants responsible for 60% of all observed power plant emissions
        - 3 Natural Gas
- Vista-CA IPCC (321)
    - 4.9% (16) of power plants responsible for 60% of all calculated power plant emissions
        - 15 Biomass, 1 Natural Gas
- CARB PMT (122)
    - 10.7% (13) of power plants responsible for 60% of all reported power plant emissions
        - 6 Biomass, 7 Natural Gas
- EPA FLIGHT (77)
    - 18.2% (14) power plants responsible for 60% of all reported power plant emissions
        - 14 Natural Gas

S4.3.6 Fugitive vs Process-Based Analysis/Determination

HyTES:

- Looked at all the time series for all 13 SPWP and CPWP with HyTES Flights/Observations
    - 9/13 are Natural Gas (2 CPWP and 7 SPWP)
    - 3/13 are Other (2 SPWP and 1 CPWP)
    - 1/13 are Coal (1 SPWP)
- More Flights = More Plumes
    - Same as AVIRIS-NG
- 9/13 were flown more than 10 times
- 4/13 were flown 4 – 8 times

AVIRIS-NG:

- Looked at all the time series for all 24 SPWP and CPWP with AVIRIS-NG Flights/Observations
    - 13/24 are Natural Gas
    - 6/24 are Biomass
    - 5/24 are Other
- Top 3/24 AVIRIS-NG power plant (also happen to be natural gas plants SPWP) emitters have been flown 1 time each and been observed to be emitting that single time at a very significantly high rate (80+kg/hr), we can at the very least say those 3 observations were fugitive or one-off emissions but are unable to confidently determine if they were persistent or consistent since repeat flights were not conducted over those three facilities. All three are natural gas SPWP.
    - PWP000051 (Natural Gas)
    - PWP000247 (Natural Gas)
    - PWP000126 (Natural Gas)
- 4/24 (16%) show persistent process-based emission behavior. 1 SPWP and 3 CPWP's have shown consistent/persistent emissions with more than 12 plus flights conducted on each facility. This is indicative of process-based emissions most likely at the component level rather than fugitive leaks since these observations are months/years apart and were still evident and persisted. At the very least, these are fugitive leaks with the potential of looking more like process-based persistent emissions. All except one have some sort of natural gas influence/component. We can target these plants for mitigation by making sure their process-based leaks are plugged.
    - SPWP
        - PWP000135 (Natural Gas) Flown 21 times, 15 plumes
    - CPWP
        - PWP000046 (Natural Gas/Other), flown 17 times, 8 plumes
        - PWP000229 (Natural Gas/Other), flown 12 times, 14 plumes
        - PWP000305 (Dairy Biogas), flown 15 times, 10 plumes
    - Based on this data, CPWP are more likely to exhibit persistent process-based behavior than SPWP.
- The other 17 SPWP's and CPWP's
    - 7 have been flown 10+ times and have only been showing 1 - 6 plumes which is more indicative of possible fugitive leaks rather than persistent process-based leaks. (The one showing 6 plumes flown 12 times is more indicative of a process-based emission than fugitive).
        - 2 Biomass (All CPWP)
        - 2 Natural Gas (All CPWP)
        - 3 Other (All CPWP)
    - 9 have been flown 4 – 9 times at most only showing between 1- 3 plumes which is more indicative of possible fugitive leaks rather than persistent process-based leaks but it is difficult to say with confidence because we

need additional flight observations in order to determine this with more certainty.

    - 2 Biomass (All CPWP)
    - 6 Natural Gas (4 SPWP and 2 CPWP)
    - 1 Other (CPWP)

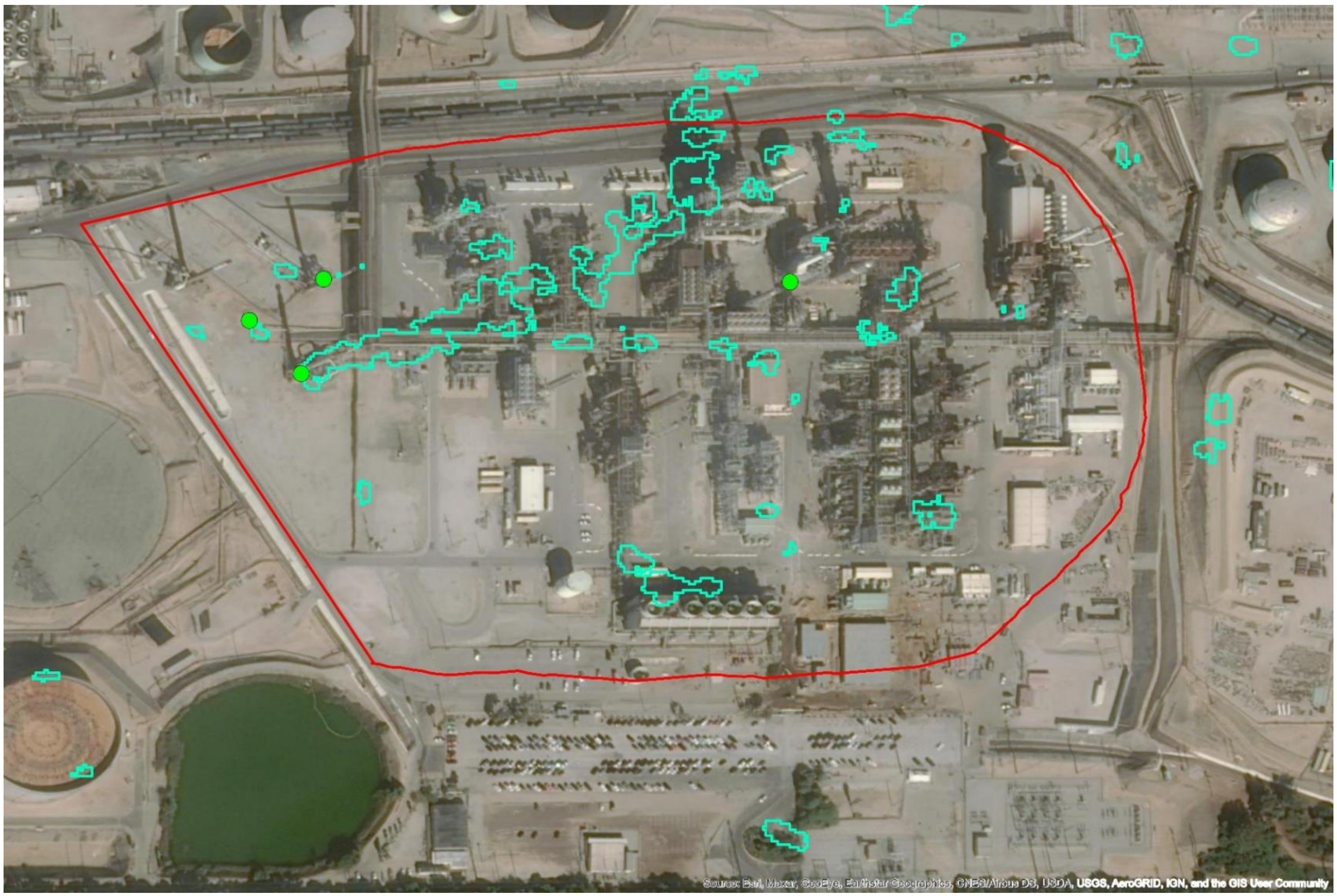
- 1 has been flown 1 time and was observed with 1 plume, which we cannot say if it is because of a persistence process-based or isolated fugitive instance with confidence
    - PWP000021, high emission estimate (52 kg hr^{-1}) Biomass Landfill gas (CPWP)

S4.3.7 Analysis of CO_2:CH_4 Ratios

Ratios for kg CO_2 a^{-1}:kg CH_4 a^{-1} were generated for CARB PMT and EPA FLIGHT facilities to determine whether a systemic relationship exists within each dataset. Currently, there is no CO_2 data available for Vista-CA or AVIRIS-NG, thus we utilized EPA's Air Markets Program Data (https://ampd.epa.gov/ampd/) that collects information from facilities via continuous emissions monitoring systems (CEMS) to generate comparable kg CO_2 a^{-1}:kg CH_4 a^{-1} ratios. Both CO_2:CH_4 ratios for CEMS to Vista-CA and CEMS to AVIRIS-NG show no correlation (Figure S4.3).[31]

Analysis of kg CO_2 a^{-1}:kg CH_4 a^{-1} ratios suggests that bottom-up CH_4 emissions from EPA FLIGHT and CARB PMT, which are described as reported, display a systemic relationship possibly due to a process-based equation or underlying assumption of CH_4 being emitted with CO_2 (Figure S4.3). Rather than collecting actual data for CH_4 emissions from reporting facilities which could be due to a wide variety of reasons, CARB and EPA could be using CO_2 data as a proxy for CH_4 emissions on a per facility basis.[20, 35] One reason for this is that fugitive emissions are potentially unaccounted as part of a power plant facility and are rather associated to the natural gas supply system.[36] Another reason is the methods used to calculate emissions from power plants. For example, EPA assumes that CH_4 is emitted as a byproduct of incomplete fuel combustion in power plants and is eventually oxidized to CO_2, which is not completely valid as our bottom-up and top-down data show.[31, 35] Both the ratio of the modeled CH_4 emissions in Vista-CA bottom-up to the EPA AMPD CEMS CO_2 data and the ratio of observed CH_4 emissions from the AVIRIS-NG source data to the CO_2 EPA AMPD CEMS data showed no systematic relationship as identified in the CARB and EPA datasets (Figure S4.3). These results suggest that most power plant CH_4 emissions are fugitive which are underestimated in these reporting programs.

Figure S4.1 Vista-CA power plant PWP000132 with HyTES plumes and AVIRIS-NG plume points identified using GSAAM. Red outlines the boundary of the power plant. Green Dots highlight centroid of AVIRIS-NG plumes.[27] Light blue polygons outline HyTES detected plumes.[28, 29, 30]



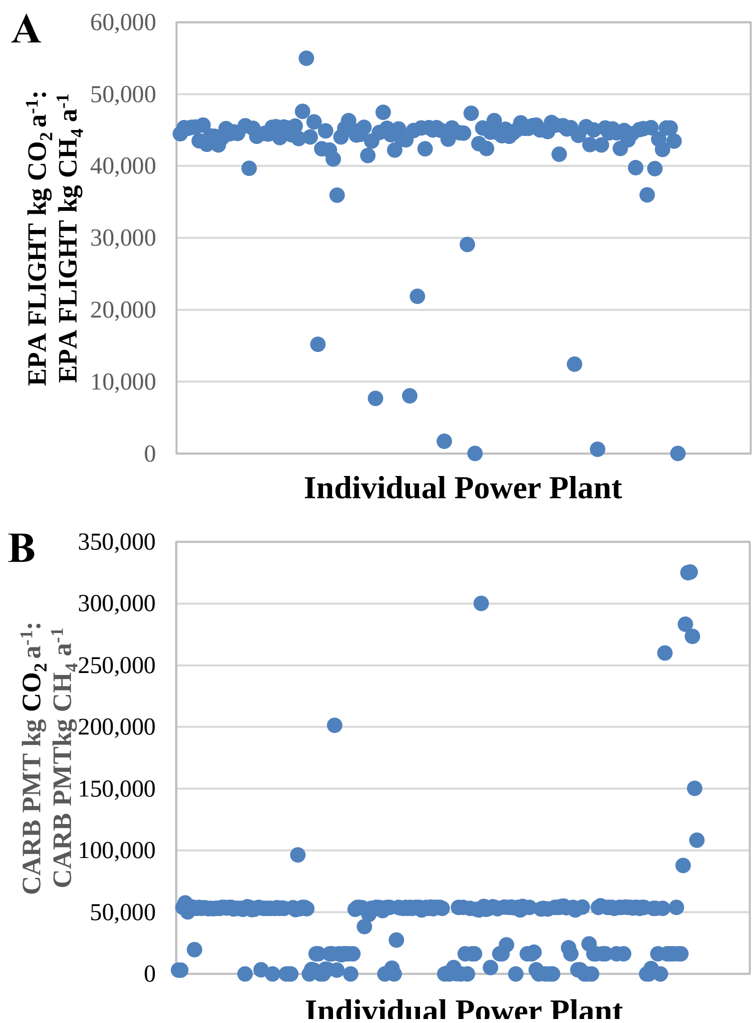


Figure S4.2 A) and B) Ratios kg CO_2 a^{-1}:kg CH_4 a^{-1} for A) EPA FLIGHT kg CO_2 a^{-1}:kg CH_4 a^{-1}, B) CARB PMT kg CO_2 a^{-1}:kg CH_4 a-1. Systemic assumptions are evident with the trends focusing on the 45,000 mark in A and 50,000 mark in B. Each power plant is listed in its own column along the horizontal axis.

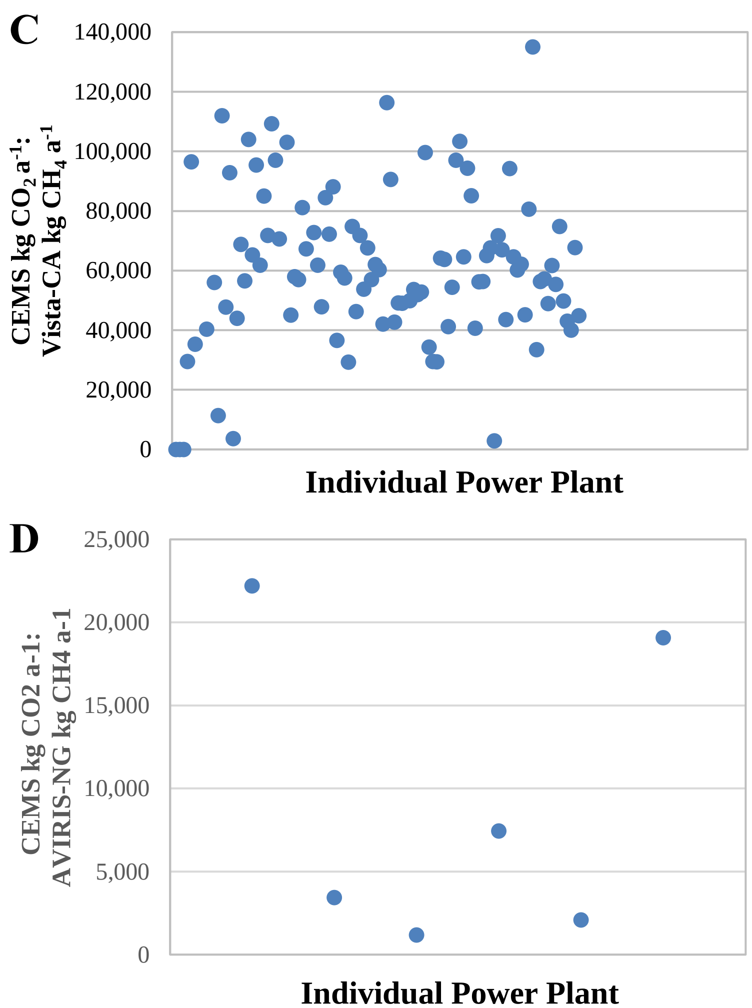


Figure S4.2 C) and D) Ratios kg CO_2 a^{-1}:kg CH_4 a^{-1} for C) CEMS kg CO_2 a^{-1}:Vista-CA Bottom-Up kg CH_4 a^{-1}. D) CEMS CO_2: AVIRIS-NG kg CH_4 a^{-1}. There is no relationship between CO_2 and CH_4 in both Vista-CA in C and AVIRIS-NG in D. Each power plant is listed in its own column along the horizontal axis.

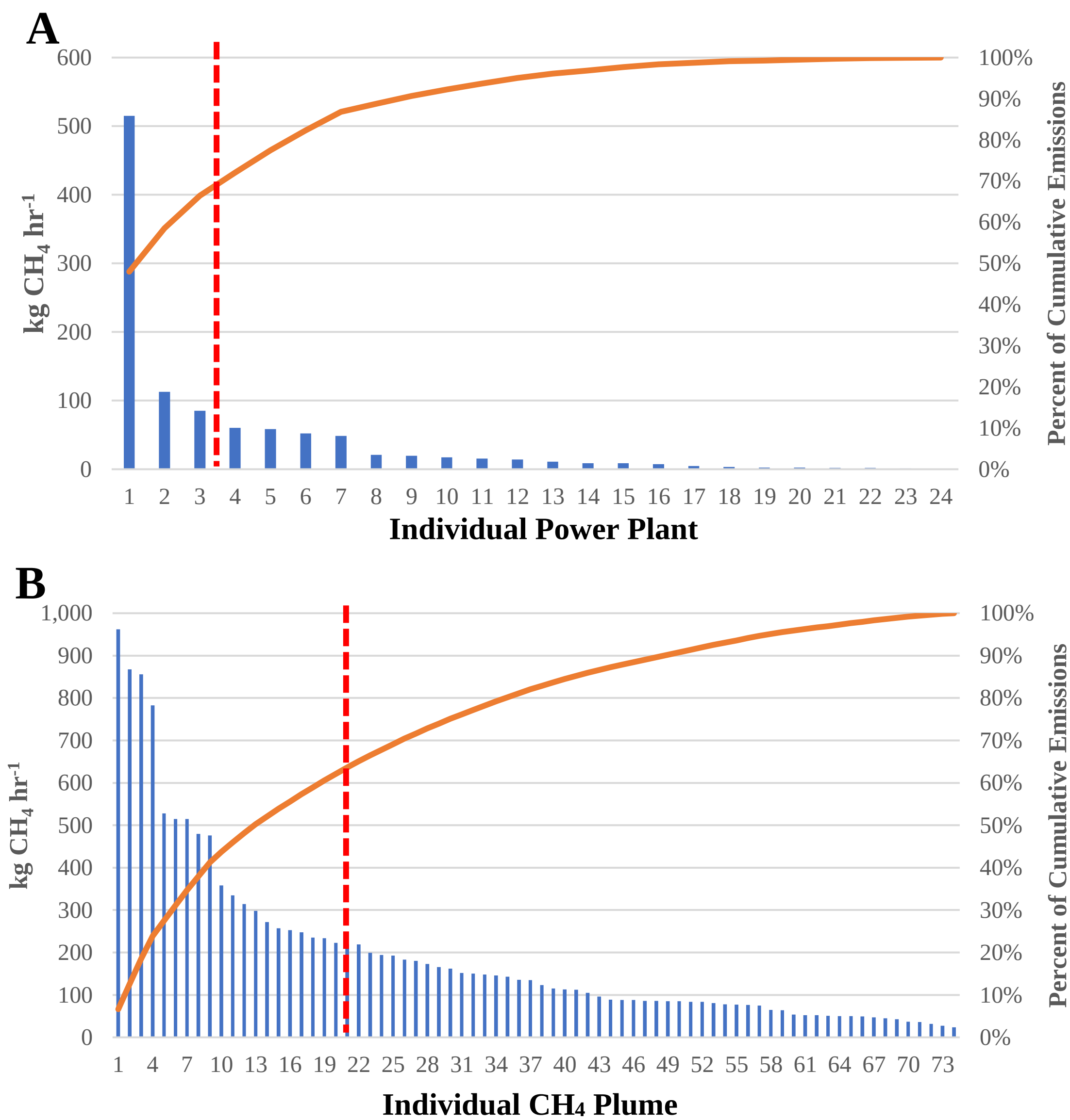


Figure S4.3 AVIRIS-NG power plant super-emitter distribution, where CH_4 emissions for the 24 power plants with 75 observed plumes are organized from largest to smallest emissions. A) shows distribution by facilities and B) shows distribution by CH_4 plumes. The orange line signifies the cumulative percentage of CH_4 emissions. The blue bars indicate the magnitude of emissions in kg CH_4 hr^{-1} for each of the matched facilities from each dataset. The red dashed line indicates the 60% super-emitter definition mark. The 60% cumulation mark is achieved with 3 SPWP natural gas facilities and 19 plumes.

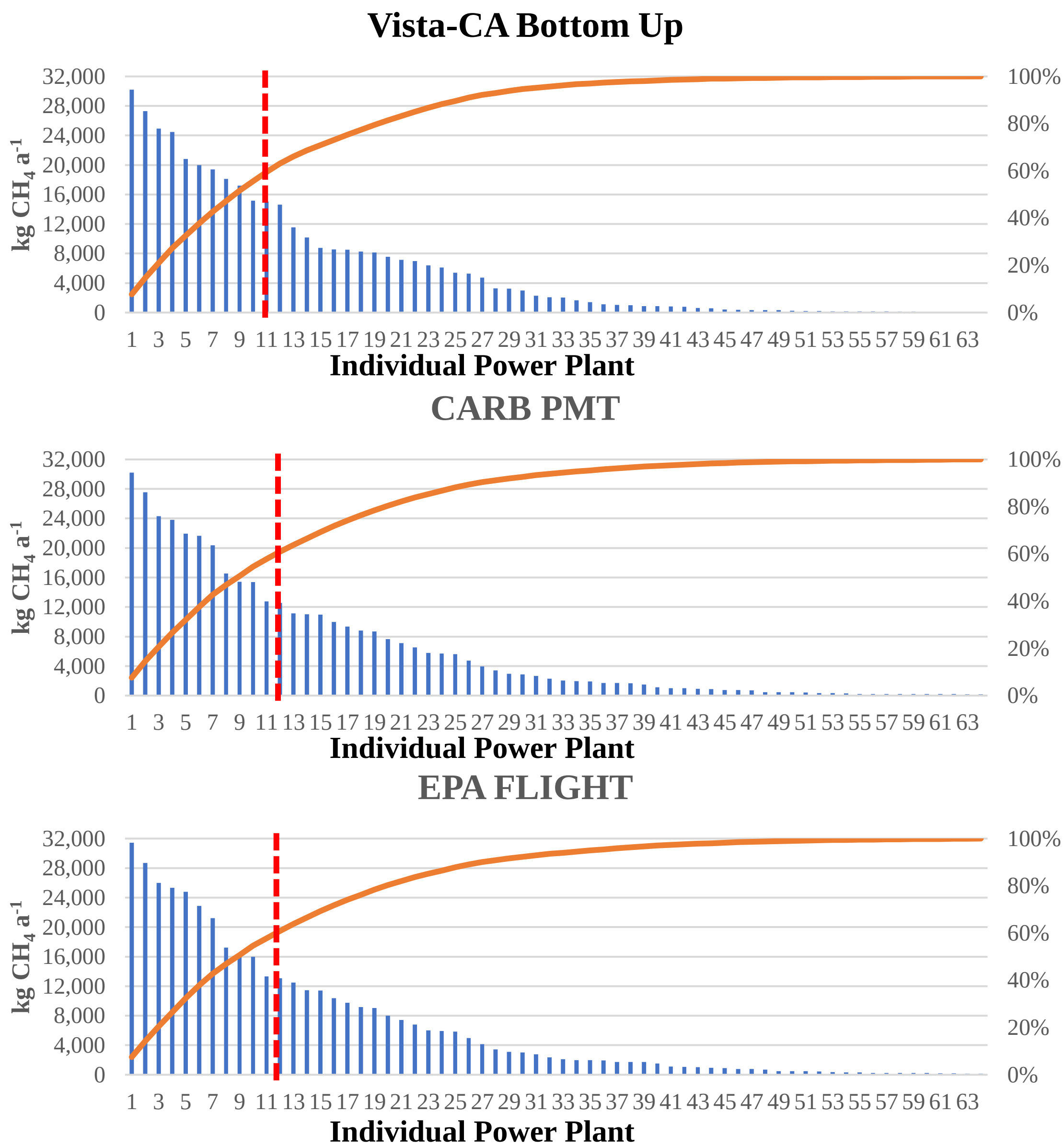


Figure S4.4 Super-emitter influence distribution comparison among the 64 power plants that are matched across all three datasets. The orange line signifies the cumulative percentage of CH_4 emissions. The blue bars indicate the magnitude of emissions in kg CH_4 a^{-1} for each of the matched facilities from each dataset. The red dashed line indicates the 60% super-emitter definition mark. For Vista-CA Bottom Up, the super-emitter mark is reached with the preceding 11 power plants. For CARB PMT, this mark is reached with the preceding 12 power plants. For EPA FLIGHT, this mark is reached with the preceding 12 power plants.

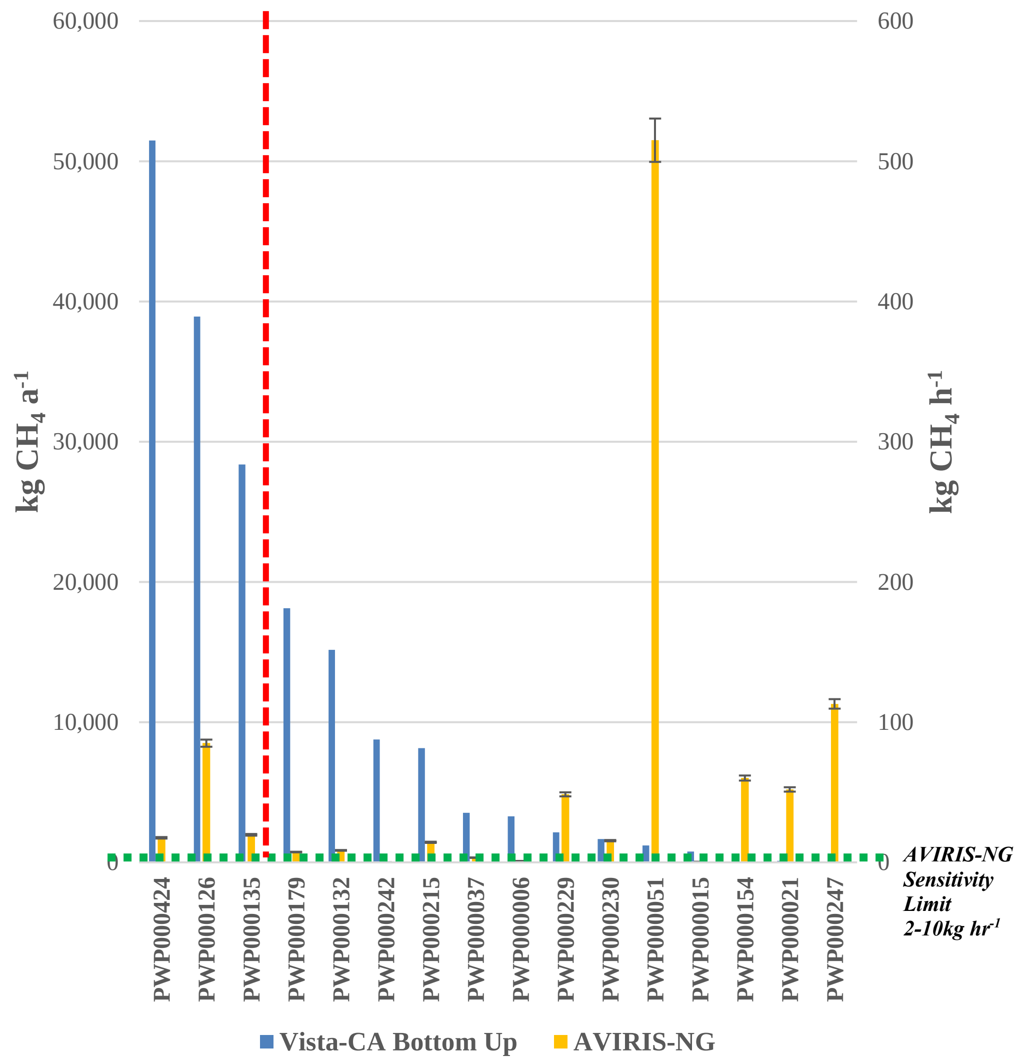
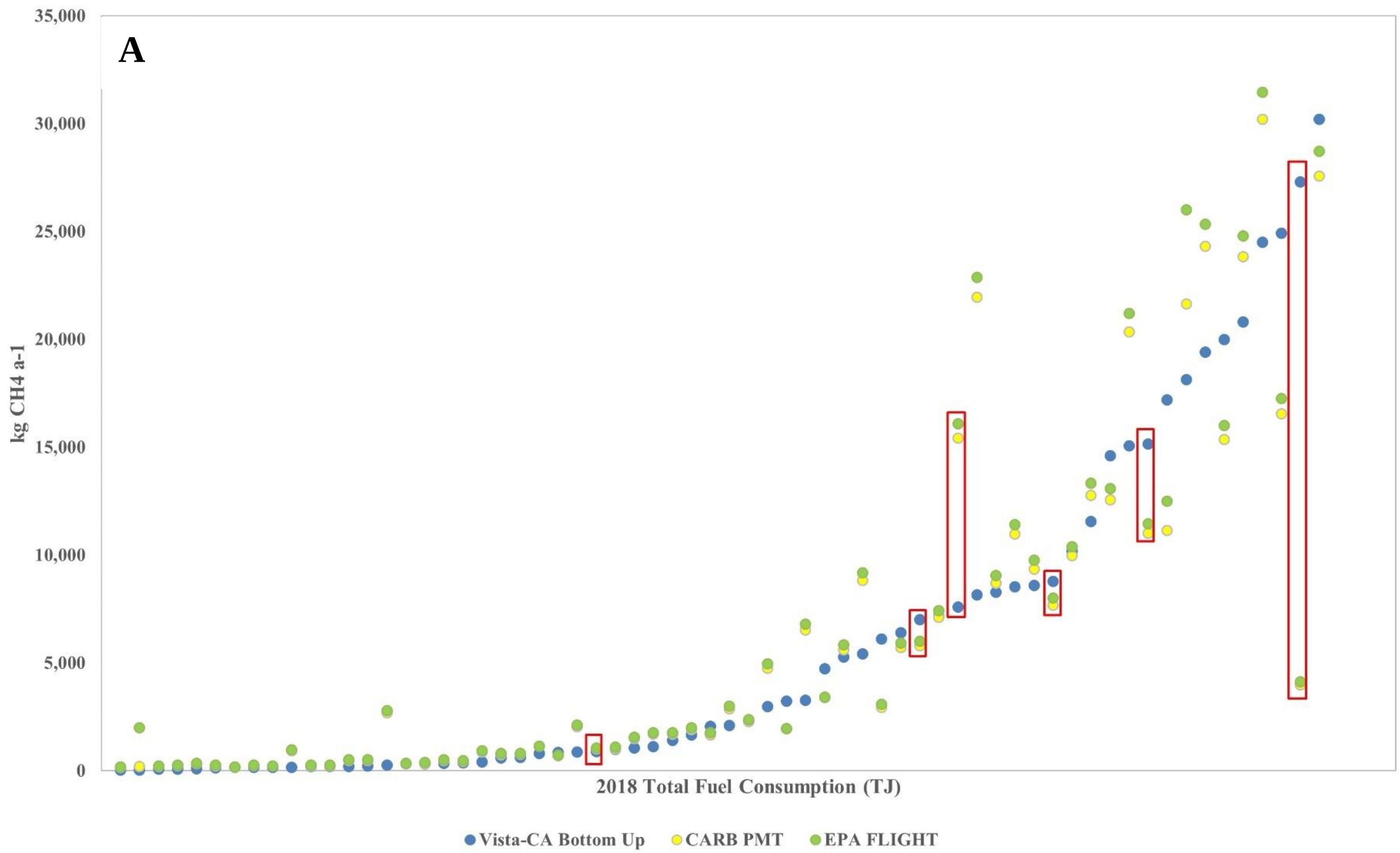


Figure S4.5 Vista-CA and AVIRIS-NG power plant super-emitter distribution comparison. The blue and yellow bars indicate the magnitude of emissions for each of the 16 matched power plant facilities from each dataset in kg CH_4 a^{-1} and kg CH_4 h^{-1} for Vista-CA (kg CH_4 a^{-1}) and AVIRIS-NG (kg CH_4 h^{-1}), respectively. The red dashed line indicates the 60% super-emitter definition mark in the Vista-CA Bottom-Up dataset. The 60% cumulation mark is achieved with 2 SPWP's and 1 CPWP in the Vista-CA emissions dataset while this mark is obtained with 1 SPWP natural gas facility in the AVIRIS-NG detected emissions. The sensitivity of the AVIRIS-NG instrument of around 2 – 10 kg h^{-1} is indicated by the green dotted bar and the error bars signify the percentage error at 3%.

Figure S4.6 A) Power plant CH_4 emission comparison for Vista-CA Bottom Up, CARB PMT, and EPA FLIGHT. Each plant contains 3 dots, representing modeled annual emissions from each data source. All plants shown in this graph were natural gas-firing. The 6 red rectangles highlight the 6 CPWP's in this distribution while the remaining 58 plants are SPWP's



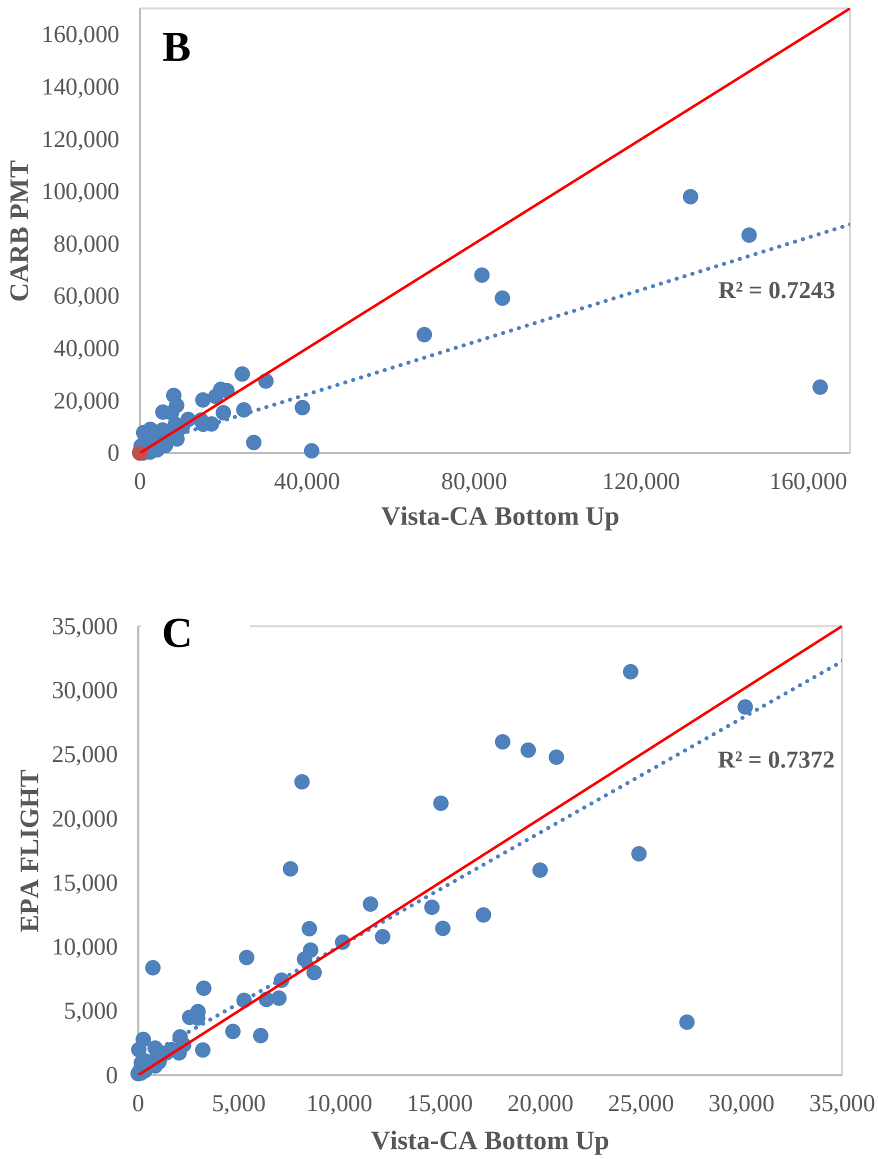


Figure S4.6 B) Vista-CA – CARB PMT power plant emission correlation (units in kg Ch_4 a^{-1}) (red line indicates the 1:1 line); C.) Vista-CA – EPA FLIGHT power plant emission correlation (units in kg CH_4 a^{-1}) (red line indicates the 1:1 line).

A

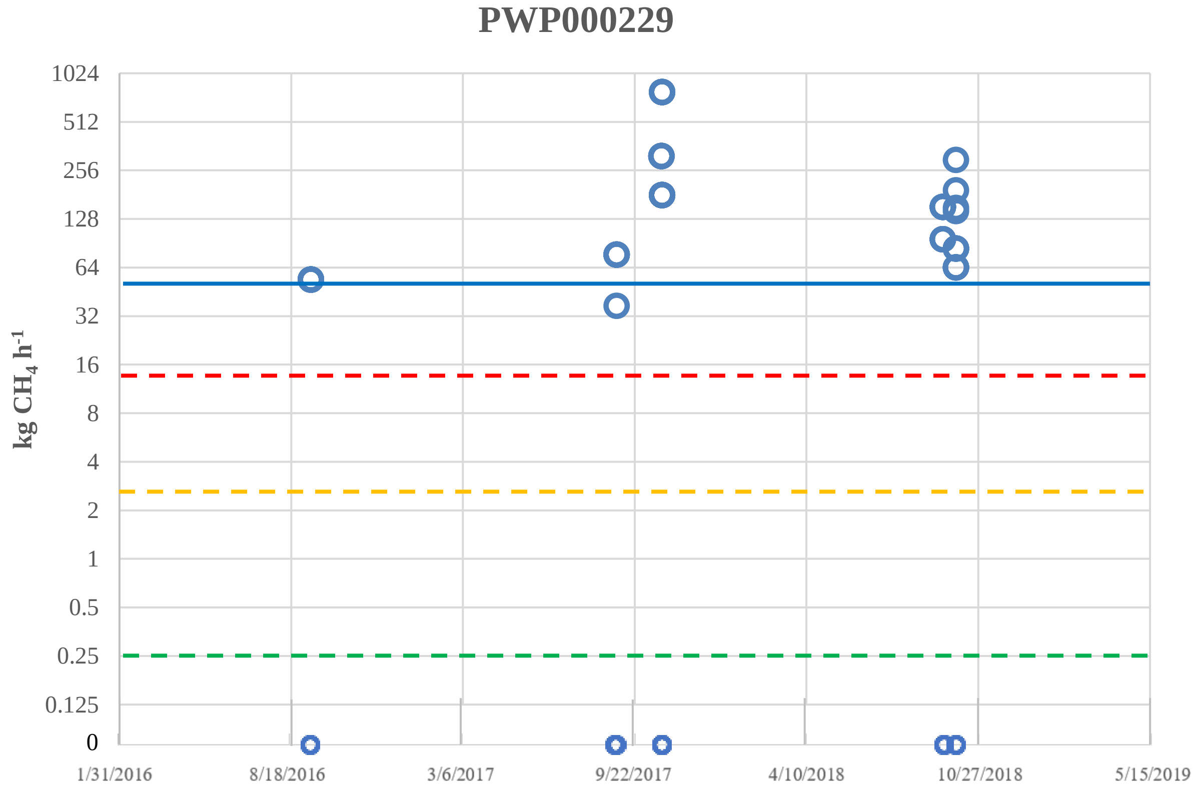


Figure S4.7 Time series plot of A.) CPWP (collocated with a refinery) Emissions are scaled logarithmically for clarity. Dashed line outline Vista-CA emissions for each power plant based on concentrated hours of assumed operation. Red dashed line indicates 175.2 hours, yellow dashed line indicates 876 hours, and green dashed line indicates 8,760 hours.

B

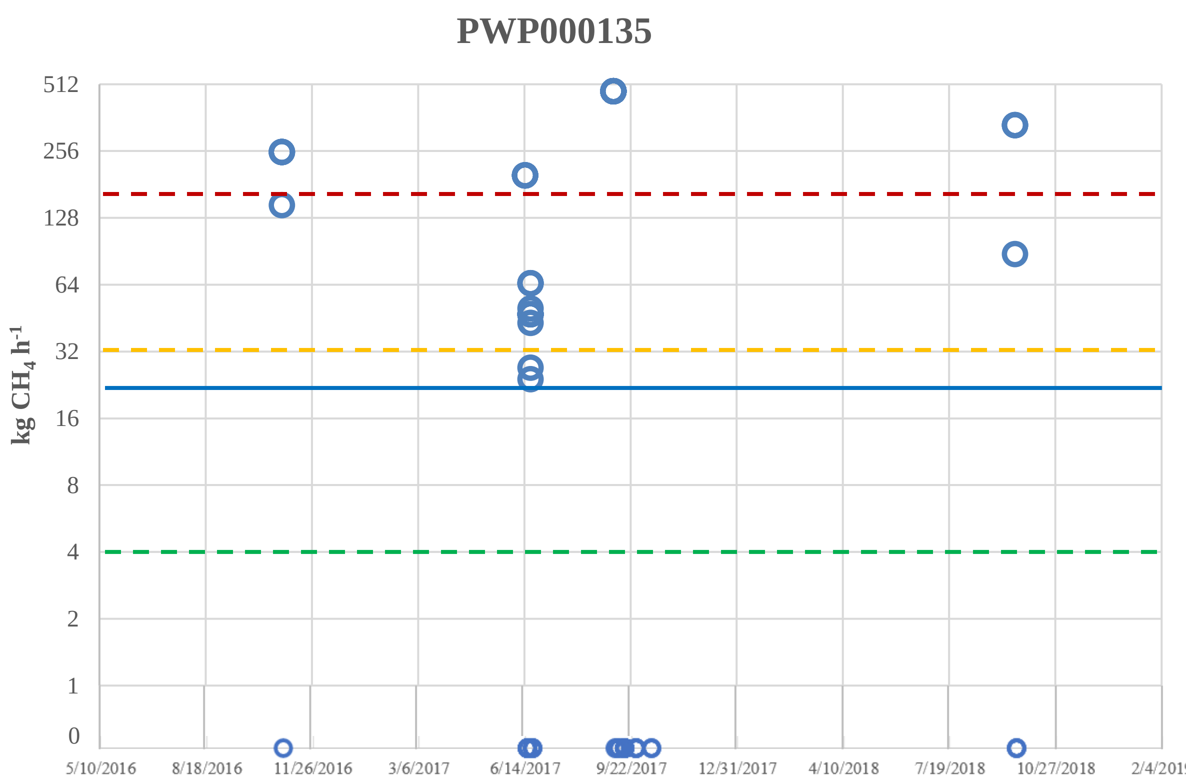


Figure S4.7 Time series plot of B.) SPWP. Emissions are scaled logarithmically for clarity. Dashed line outline Vista-CA emissions for each power plant based on concentrated hours of assumed operation. Red dashed line indicates 175.2 hours, yellow dashed line indicates 876 hours, and green dashed line indicates 8,760 hours.

Table S4.1 A.) AVIRIS-NG attributed observation breakdown by Vista-CA power plant. This table summarizes the power plants flown by AVIRIS-NG that were detected as emitting CH_4. The tables detail the number of flights conducted over a given plant, the total number of plumes found along with sources identified, and the location and description of those sources.

A

Vista-ID	AVIRIS-NG (kg CH4 h-1)	Vista-CA (kg CH4 a-1)	Total AVIRIS-NG Flights	Sources (Number of locations)	Total AVIRIS-NG Attributed Plumes	Type	Persistency	Emission Type	Location 1	Location 2	Location 3	Location 4	Location 5	Location 6
PWP000006	1.0	3272.9	9	1	1	SPWP	one-time	fugitive	Electric Switchyard 1					
PWP000015	0.4	766.7	11	1	1	CPWP	one-time	fugitive	Building 1					
PWP000021	52.0	117.6	1	1	1	CPWP	one-time	fugitive	Refuse Pile 1					
PWP000037	3.3	3536.6	8	2	2	CPWP	persistent 2x	fugitive	Storage Tank 1	Storage Tank 1				
PWP000046	11.0		17	4	8	CPWP	persistent 3x	process-based	Infrastructure 1	Open Field 3	Open Field 2	Infrastructure 2		
PWP000051	515.0	1194.5	1	1	1	SPWP	one-time	fugitive	Building 1					
PWP000121	4.6		14	2	2	CPWP	one-time	process-based	Storage Tank 1	Open Field 1				
PWP000126	85.0	38912.8	1	1	1	SPWP	one-time	process-based	Infrastructure/Stack 1					
PWP000132	8.5	15152.7	12	4	6	CPWP	persient 5x	process-based	Infrastructure 1	Infrastructure 1	Infrastructure 2	Infrastructure 2		
PWP000135	19.6	28370.8	21	1	1	SPWP	one-time	process-based	Infrastructure 1					
PWP000145	1.9		15	1	3	CPWP	persistent 3x	fugitive	Infrastructure 3					
PWP000154	60.1	141.4	4	1	1	CPWP	one-time	fugitive	Open Field 1					
PWP000179	7.3	18131.0	7	1	2	SPWP	persistent 2x	fugitive	Infrastructure/Stack 2					
PWP000215	14.3	8148.6	6	1	1	SPWP	one-time	fugitive	NG infrastructure 1					
PWP000229	48.5	2145.4	12	6	14	CPWP	persistent 3x	fugitive	Storage Tank 1	Storage Tank 2	Storage Tank 6	Infrastructure 1	Infrastructure 3	Infrastructure 1
PWP000230	15.5	1664.0	8	1	3	SPWP	persistent 3x	fugitive	Infrastructure 3					
PWP000242	0.1	8766.4	7	1	1	CPWP	one-time	process-based	Infrastructure/Stack 1					
PWP000247	113.0	23.9	1	1	1	SPWP	one-time	fugitive	Open Field 1					
PWP000285	8.5		13	1	3	CPWP	persistent 3x	process-based	Open Field 3					
PWP000305	58.4		15	2	10	CPWP	persistent 5x	fugitive	Lagoon 3	Lagoon 7				
PWP000323	2.3		7	2	2	CPWP	persistent 2x	fugitive	Infrastructure 1	Infrastructure 1				
PWP000388	21.1		5	1	1	CPWP	one-time	fugitive	Infrastructure 1					
PWP000420	2.3		14	1	2	CPWP	persistent 2x	fugitive	Infrastructure/Road 2					
PWP000424	17.5	51492.4	13	1	3	CPWP	persistent 3x	fugitive	Storage Tank 3					

B

Vista-ID	Fuel Source	Total HyTES Attributed Plumes	Total HyTES Flights	Type
PWP000029	Coal	35	35	SPWP
PWP000033	Natural Gas	3	12	SPWP
PWP000046	Other	13	13	SPWP
PWP000046	Other	17	17	CPWP
PWP000121	Natural Gas	3	3	CPWP
PWP000121	Natural Gas	11	11	SPWP
PWP000132	Natural Gas	45	45	CPWP
PWP000133	Natural Gas	35	101	SPWP
PWP000141	Other	4	4	CPWP
PWP000178	Natural Gas	5	5	CPWP
PWP000207	Natural Gas	2	8	SPWP
PWP000285	Natural Gas	3	11	SPWP
PWP000285	Natural Gas	66	256	CPWP
PWP000414	Natural Gas	7	30	SPWP
PWP000420	Other	16	16	CPWP
PWP000454	Natural Gas	1	4	SPWP

Table S4.1 B.) HyTES attributed observation breakdown by Vista-CA power plant. Th**is** table summarizes the power plants flown by HyTES that were detected as emitting CH_4. The tables detail the number of flights conducted over a given plant, the total number of plumes found along with sources identified, and the location and description of those sources

Table S4.2 Power plants observed by AVIRIS-NG with no emissions. This table provides a list of all Vista-CA power plants that were flown by AVIRIS-NG but were not observed to be emitting any CH_4 at the time of observation. For each power plant, coordinates are provided, the power plant type, and the total number of flights AVIRIS-NG conducted over that power plant.

Vista-ID	Latitude	Longitude	Power Plant Type	Total Number of AVIRIS-NG Flights Flown
PWP000001	35.16598177	-119.1065063	SPWP	41
PWP000003	35.95333604	-120.8591409	Oil and Gas Facility	2
PWP000004	33.76919138	-118.100978	SPWP	7
PWP000005	33.64444727	-117.978604	SPWP	8
PWP000008	34.04156467	-117.3605156	SPWP	1
PWP000012	37.57466407	-120.9852294	SPWP	1
PWP000014	37.73257081	-121.1162515	SPWP	3
PWP000016	34.08840491	-117.2460224	SPWP	2
PWP000019	34.43174678	-118.6428042	Landfill	8
PWP000020	37.88328896	-121.1851666	Landfill	6
PWP000022	38.00385809	-121.934572	Landfill	19
PWP000023	37.50428975	-122.4060061	Landfill	11
PWP000026	37.75474775	-121.7289581	Landfill	17
PWP000027	37.10679609	-120.248717	SPWP	1
PWP000030	33.68859501	-117.8362438	SPWP	3
PWP000031	35.4837382	-119.0298414	SPWP	8
PWP000035	35.08989853	-119.4408423	SPWP	2
PWP000036	35.09247043	-119.4440113	SPWP	2
PWP000039	33.90696733	-118.0130896	SPWP	1
PWP000040	33.93306841	-117.8555642	Processing Plant	6
PWP000045	33.71859723	-117.7097901	Landfill	9
PWP000047	33.93331675	-117.841892	Landfill	7
PWP000052	38.05704044	-122.219354	SPWP	3
PWP000053	34.15108169	-118.7237674	Landfill	5
PWP000054	36.29519703	-119.4135159	Digester	2
PWP000055	34.138467	-118.1256	SPWP	1
PWP000062	34.1358	-118.1267	SPWP	1
PWP000065	33.87595065	-118.2493429	SPWP	3
PWP000067	34.14484951	-118.390888	SPWP	1
PWP000068	37.99785803	-122.0678511	SPWP	2
PWP000071	33.92881392	-118.1056538	SPWP	2
PWP000072	33.94399661	-118.4042446	SPWP	8
PWP000073	37.71941136	-120.8986107	Digester	4

PWP000075	34.38065567	-118.5001697	SPWP	9
PWP000076	35.09685309	-119.4298422	SPWP	3
PWP000078	37.10728307	-120.2484327	SPWP	1
PWP000081	36.18340393	-119.3706909	WWTP	6
PWP000083	33.89182264	-117.609994	WWTP	2
PWP000085	36.15546196	-120.3970692	Oil and Gas Facility	2
PWP000086	36.18063357	-120.3883361	Oil and Gas Facility	2
PWP000087	36.19043147	-120.356759	SPWP	2
PWP000088	36.23737202	-120.3687594	Oil and Gas Facility	2
PWP000090	35.18792459	-120.5936678	Landfill	1
PWP000092	39.3657123	-122.2667333	SPWP	1
PWP000093	38.08817785	-121.3867402	WWTP	4
PWP000094	33.99476406	-118.1536306	SPWP	2
PWP000095	33.88182147	-117.5562357	SPWP	1
PWP000102	37.38537357	-121.1405388	Landfill	1
PWP000104	38.05694827	-122.2161522	SPWP	4
PWP000108	33.88008675	-117.886809	SPWP	2
PWP000109	32.79617582	-116.9720526	SPWP	1
PWP000110	35.36382092	-119.659227	SPWP	7
PWP000111	35.36363747	-119.6753619	Oil and Gas Facility	3
PWP000112	35.346222	-119.643367	SPWP	3
PWP000115	38.0176174	-121.8456398	WWTP	8
PWP000120	35.22080301	-119.5833602	SPWP	4
PWP000123	35.50064167	-119.0460671	SPWP	8
PWP000124	34.06059843	-117.3536386	SPWP	2
PWP000129	32.79716174	-116.9717	SPWP	1
PWP000130	32.80251713	-115.5399568	SPWP	1
PWP000131	37.18686141	-120.4898912	SPWP	2
PWP000133	33.91051879	-118.4248717	SPWP	9
PWP000134	35.28067457	-119.4726655	Oil and Gas Facility	21
PWP000136	34.43131338	-119.8999995	SPWP	2
PWP000138	33.13633825	-117.3365974	SPWP	2
PWP000141	33.78891903	-118.2350521	Refinery	17
PWP000144	34.09137009	-117.5273129	SPWP	3
PWP000146	34.48368246	-120.042768	Processing Plant	3
PWP000149	38.02054544	-122.0668446	Refinery	8
PWP000151	36.61720142	-120.0993512	SPWP	2
PWP000157	38.01640767	-121.7585409	SPWP	2
PWP000158	34.4753001	-120.2052162	Oil and Gas Facility	2
PWP000164	37.40122666	-121.9687148	SPWP	4

PWP000165	36.99991773	-121.5365424	SPWP	3
PWP000166	36.99991773	-121.5365424	SPWP	3
PWP000167	34.1264471	-118.1493235	SPWP	3
PWP000169	38.225635	-121.8410688	SPWP	1
PWP000170	34.08837591	-117.5336764	SPWP	2
PWP000171	34.15533313	-118.2783729	SPWP	3
PWP000174	33.61307454	-117.8219577	SPWP	3
PWP000175	33.99880407	-118.2210221	SPWP	1
PWP000176	36.27186138	-119.6493912	SPWP	2
PWP000177	33.76985496	-118.2654672	SPWP	12
PWP000178	33.78027449	-118.240243	SPWP	15
PWP000184	35.51563618	-119.0401032	SPWP	8
PWP000187	33.62279418	-117.9363657	SPWP	6
PWP000188	33.838056	-118.315	SPWP	12
PWP000192	37.91187939	-121.261534	SPWP	2
PWP000193	33.73848806	-117.1694656	SPWP	2
PWP000194	37.38505206	-121.9637212	SPWP	3
PWP000196	37.80767675	-121.2776377	SPWP	1
PWP000199	37.40834129	-122.0292213	SPWP	1
PWP000200	33.91958435	-118.1277963	SPWP	1
PWP000201	34.03201335	-117.608654	SPWP	1
PWP000202	36.32784195	-119.2950532	SPWP	1
PWP000205	35.50907382	-119.0305462	SPWP	8
PWP000206	35.29478185	-118.917208	Refinery	5
PWP000207	35.45217504	-118.9850804	SPWP	8
PWP000208	35.44066983	-118.9612838	Oil and Gas Facility	7
PWP000209	38.51478672	-121.1925225	Landfill	6
PWP000213	36.53936956	-119.5800086	SPWP	1
PWP000216	34.177215	-118.3148342	SPWP	1
PWP000220	33.13879807	-117.2859613	SPWP	2
PWP000222	33.79425401	-118.2344901	Refinery	17
PWP000223	35.49619334	-119.0088773	SPWP	8
PWP000224	37.41324482	-122.0275584	SPWP	2
PWP000226	38.08819	-121.3876678	WWTP	4
PWP000228	33.7643334	-118.2245818	SPWP	11
PWP000231	38.03011767	-121.8730424	SPWP	6
PWP000232	35.66570458	-119.76715	SPWP	4
PWP000233	34.1784842	-118.3145726	SPWP	1
PWP000234	36.68981123	-119.7404372	SPWP	1
PWP000235	33.99880389	-118.2210404	SPWP	1

PWP000238	34.20664618	-119.2508625	SPWP	1
PWP000239	36.71366151	-121.7679346	Landfill	5
PWP000240	37.78925525	-121.6016694	SPWP	3
PWP000241	38.01695601	-121.7651433	SPWP	2
PWP000243	38.03382935	-122.1147804	SPWP	6
PWP000244	37.40510421	-121.9496625	SPWP	10
PWP000248	34.20559028	-119.2477508	SPWP	1
PWP000249	35.3159958	-119.659614	SPWP	10
PWP000250	35.31950142	-119.6619932	SPWP	11
PWP000251	32.90404402	-115.5127812	SPWP	1
PWP000252	32.90690576	-115.5131102	SPWP	1
PWP000253	37.22084792	-121.7463284	SPWP	1
PWP000254	34.14251679	-117.4273534	Landfill	1
PWP000255	35.19423877	-119.5709163	SPWP	3
PWP000259	33.69102641	-117.8332073	SPWP	3
PWP000260	34.0065835	-117.5608291	SPWP	1
PWP000261	34.00797005	-117.5638959	SPWP	1
PWP000263	34.29275664	-118.3898236	Landfill	1
PWP000266	34.0336783	-117.9063581	Landfill	3
PWP000267	38.59633008	-121.6874324	Landfill	1
PWP000268	36.70610913	-121.769885	SPWP	5
PWP000269	34.08192792	-117.2419178	SPWP	2
PWP000274	32.72624134	-117.1464282	SPWP	1
PWP000276	34.04539506	-117.5404486	SPWP	2
PWP000281	35.27621644	-119.5987218	Oil and Gas Facility	8
PWP000282	37.96764791	-122.3820398	SPWP	2
PWP000283	32.71489	-117.1689257	SPWP	1
PWP000286	33.93272724	-117.841073	Landfill	7
PWP000287	34.17657158	-118.314606	SPWP	1
PWP000288	34.32613964	-118.4457053	SPWP	2
PWP000289	33.98962307	-117.6807993	SPWP	5
PWP000290	37.79351982	-122.3944033	SPWP	1
PWP000295	34.12922477	-119.168989	SPWP	1
PWP000299	32.57360796	-116.9129016	SPWP	1
PWP000301	34.14252422	-119.1830268	SPWP	1
PWP000302	34.14265913	-119.1853777	WWTP	1
PWP000311	33.8985072	-118.1465443	Refinery	3
PWP000312	37.76123569	-122.4562531	SPWP	2
PWP000313	34.9558962	-118.8438908	SPWP	5
PWP000316	38.01935915	-122.2372043	SPWP	8

PWP000317	38.04171931	-122.2587992	Refinery	16
PWP000318	32.57384457	-116.9179962	SPWP	1
PWP000319	34.46375025	-118.5940217	SPWP	7
PWP000320	38.03965331	-121.8947702	SPWP	11
PWP000321	37.83294563	-122.2840369	SPWP	1
PWP000322	33.69225904	-117.9385917	WWTP	3
PWP000325	38.21096121	-121.9756483	Landfill	13
PWP000326	37.82507966	-122.2949887	WWTP	2
PWP000328	33.4948786	-117.6145805	Landfill	2
PWP000329	34.02380591	-118.0244314	SPWP	4
PWP000334	40.50893206	-122.4243657	SPWP	1
PWP000336	32.58465953	-116.9334152	SPWP	1
PWP000337	37.94220567	-122.3907159	Refinery	14
PWP000338	36.68760204	-119.7238598	SPWP	1
PWP000341	37.73151447	-121.1157188	SPWP	3
PWP000342	33.96329657	-117.4529501	WWTP	4
PWP000343	33.96218265	-117.4563154	WWTP	4
PWP000344	38.0145574	-121.7902232	SPWP	2
PWP000345	32.95554941	-115.5362425	SPWP	1
PWP000353	34.02838792	-117.6016931	WWTP	1
PWP000355	36.83536987	-119.7657915	SPWP	1
PWP000356	34.03061227	-118.4790875	SPWP	6
PWP000357	35.95172421	-120.867681	Oil and Gas Facility	2
PWP000364	32.77510993	-117.0722459	SPWP	1
PWP000369	34.95032041	-120.4139746	SPWP	1
PWP000372	35.93590589	-120.8407364	SPWP	2
PWP000374	38.53103907	-121.4003776	SPWP	1
PWP000375	33.91892022	-118.4277905	WWTP	15
PWP000379	37.73963536	-122.392852	SPWP	1
PWP000390	37.43425287	-121.9463887	WWTP	16
PWP000393	38.301503	-122.7477293	Landfill	1
PWP000394	38.30153643	-122.7476187	Landfill	1
PWP000395	38.30155854	-122.7478301	Landfill	1
PWP000398	35.44906859	-119.7336306	SPWP	3
PWP000399	35.42062374	-118.9644113	Oil and Gas Facility	7
PWP000400	33.75948422	-118.2403321	SPWP	9
PWP000402	34.04092588	-117.8216762	Landfill	1
PWP000404	32.91015415	-115.5677162	SPWP	2
PWP000407	34.09374347	-117.5510332	SPWP	2
PWP000409	38.36920173	-122.7655283	WWTP	1

PWP000410	37.41930675	-122.0164343	WWTP	1
PWP000411	35.2105525	-119.583845	SPWP	3
PWP000412	34.33553923	-118.5191747	Landfill	49
PWP000414	35.45490271	-119.0023395	SPWP	8
PWP000416	35.11568183	-119.4755737	Oil and Gas Facility	2
PWP000417	34.47718375	-120.1284169	Landfill	3
PWP000421	37.71993038	-121.4885787	SPWP	4
PWP000422	35.18609114	-120.6225691	Oil and Gas Facility	1
PWP000423	33.76838208	-118.2141644	SPWP	13
PWP000425	34.1462098	-118.3050836	Landfill	1
PWP000426	37.71113988	-121.4918966	SPWP	3
PWP000427	37.7951745	-122.4027941	SPWP	1
PWP000428	33.64795032	-117.8465695	SPWP	4
PWP000429	33.66541927	-117.8538333	WWTP	2
PWP000433	37.77808905	-122.4505274	SPWP	1
PWP000436	37.43291853	-121.9420547	WWTP	15
PWP000438	38.07393312	-122.1409403	Refinery	5
PWP000439	34.2444572	-118.3924221	SPWP	2
PWP000440	33.85087334	-118.3040667	SPWP	12
PWP000441	33.99883579	-118.2210433	SPWP	1
PWP000442	36.1822643	-119.6779722	Digester	7
PWP000443	36.68689381	-119.720451	SPWP	1
PWP000447	37.49000085	-120.9045045	SPWP	1
PWP000449	37.48736482	-120.8958021	SPWP	1
PWP000451	33.81645336	-118.2442377	Refinery	17
PWP000452	35.35458045	-119.6618228	SPWP	7
PWP000454	35.44075378	-119.0129252	SPWP	8
PWP000455	33.92475584	-118.0678113	SPWP	2
PWP000457	33.9706	-118.0472	SPWP	2
PWP000458	33.7789072	-118.2341079	Refinery	10
PWP000460	37.65261119	-121.0197748	SPWP	1
PWP000461	38.68879537	-121.7374604	SPWP	1
PWP000462	37.82507966	-122.2949887	WWTP	2
PWP000464	37.41569896	-122.0256478	SPWP	2
PWP000468	37.43232419	-121.955533	Landfill	16

Table S4.3 Vista-CA Level 2 Power Plants CH_4 Emissions Dataset. This table provides the emissions for Vista-CA Bottom Up, CARB PMT, EPA FLIGHT, and AVIRIS-NG along with the number of plumes that were attributed to HyTES observations for each power plant in the Vista-CA dataset.

Vista ID	Type	Fuel Source	Vista CA kg CH_4 a^{-1}	CARB PMT kg CH_4 a^{-1}	EPA FLIGHT kg CH_4 a^{-1}	AVIRIS-NG kg CH_4 h^{-1}	Number of plumes observed by HyTES (via GSAAM)
PWP000001	SPWP	biomass	5434.30	0	0	0	2
PWP000002	SPWP	natural gas	78.75	0	0	0	0
PWP000003	OGF	natural gas	730.37	0	8375	0	0
PWP000004	SPWP	natural gas	8516.52	10970	11416.67	0	0
PWP000005	SPWP	natural gas	2978.13	4750	4958.33	0	0
PWP000006	SPWP	natural gas	3272.93	6530	6791.67	1	0
PWP000007	SPWP	natural gas	60.67	0	0	0	0
PWP000008	SPWP	natural gas	4.74	0	125	0	0
PWP000010	SPWP	natural gas	128.33	0	0	0	0
PWP000011	SPWP	natural gas	585.32	740	791.67	0	0
PWP000012	SPWP	natural gas	3226.92	1950	1958.33	0	0
PWP000013	SPWP	natural gas	0.00	0	0	0	0
PWP000014	SPWP	natural gas	58.86	190	208.33	0	0
PWP000015	Landfill	biomass	766.74	2450	0	1	0
PWP000016	SPWP	natural gas	59.06	0	0	0	0
PWP000018	Landfill	biomass	34.73	0	0	0	0
PWP000019	Landfill	biomass	643.23	0	0	0	0
PWP000020	Landfill	biomass	334.00	1050	0	0	0
PWP000021	Landfill	biomass	117.63	0	0	52	0
PWP000022	Landfill	biomass	322.19	980	0	0	0
PWP000023	Landfill	biomass	938.62	0	0	0	0
PWP000024	Landfill	biomass	321.47	0	0	0	0
PWP000025	SPWP	biomass	241.00	570	0	0	0
PWP000026	Landfill	biomass	346.48	1030	0	0	0
PWP000027	SPWP	biomass	44976.76	0	0	0	0
PWP000028	SPWP	natural gas	222.05	0	833.33	0	0
PWP000029	SPWP	coal	12552.60	0	0	0	1
PWP000030	SPWP	natural gas	547.10	0	916.67	0	0
PWP000031	SPWP	natural gas	142.46	150	166.67	0	0
PWP000032	SPWP	natural gas	206.24	330	0	0	0
PWP000033	SPWP	natural gas	322.78	320	333.33	0	1
PWP000034	SPWP	natural gas	8.00	0	0	0	0
PWP000035	SPWP	natural gas	4070.46	0	0	0	0

PWP000036	SPWP	natural gas	1626.65	0	0	0	0
PWP000037	OGF	natural gas	3536.55	0	0	3	0
PWP000038	SPWP	biomass	0.00	0	0	0	0
PWP000039	SPWP	natural gas	158.53	0	0	0	0
PWP000040	Processing Plant	other	0.00	0	0	0	0
PWP000041	SPWP	biomass	0.00	0	0	0	0
PWP000042	SPWP	natural gas	11555.52	12770	13333.33	0	0
PWP000043	SPWP	other	0.00	0	0	0	0
PWP000045	Landfill	biomass	1736.80	4850	0	0	0
PWP000046	Refinery	other	0.00	0	0	11	1
PWP000047	Landfill	biomass	2460.69	0	0	0	0
PWP000048	SPWP	biomass	0.00	0	0	0	0
PWP000049	SPWP	biomass	131827.63	98030	0	0	0
PWP000050	SPWP	biomass	0.00	0	0	0	0
PWP000051	SPWP	natural gas	1194.51	0	0	515	0
PWP000052	SPWP	other	0.00	0	0	0	0
PWP000053	Landfill	biomass	511.37	1860	0	0	0
PWP000054	Digester	biomass	0.00	0	0	0	0
PWP000055	SPWP	natural gas	1050.73	0	1166.67	0	0
PWP000056	SPWP	natural gas	90.86	0	0	0	0
PWP000058	SPWP	natural gas	92.85	140	0	0	0
PWP000059	SPWP	natural gas	64.40	130	0	0	0
PWP000060	SPWP	natural gas	80.23	170	0	0	0
PWP000061	SPWP	natural gas	92.08	100	0	0	0
PWP000062	SPWP	biomass	58.09	0	0	0	0
PWP000063	SPWP	other	0.00	0	0	0	0
PWP000064	SPWP	natural gas	1403.24	1690	1750	0	0
PWP000065	SPWP	natural gas	0.00	0	0	0	2
PWP000066	WWTP	natural gas	6098.23	2940	3083.33	0	0
PWP000067	SPWP	natural gas	83.62	0	0	0	0
PWP000068	SPWP	natural gas	358.13	0	0	0	0
PWP000071	SPWP	natural gas	134.92	250	0	0	0
PWP000072	SPWP	natural gas	474.66	0	0	0	0
PWP000073	Digester	biomass	0.00	0	0	0	0
PWP000074	SPWP	natural gas	7.91	0	0	0	0
PWP000075	SPWP	natural gas	0.00	0	0	0	0
PWP000076	SPWP	natural gas	99.45	200	0	0	0
PWP000077	SPWP	petroleum	457.43	0	0	0	0
PWP000078	SPWP	natural gas	137.14	0	0	0	0
PWP000079	SPWP	natural gas	67.71	70	0	0	0

PWP000080	Landfill	natural gas	0.12	0	0	0	0
PWP000081	WWTP	biomass	0.00	0	0	0	0
PWP000082	SPWP	natural gas	1159.91	0	0	0	0
PWP000083	WWTP	natural gas	79.30	230	0	0	0
PWP000084	SPWP	natural gas	85.44	0	0	0	0
PWP000085	OGF	natural gas	1156.09	0	0	0	0
PWP000086	OGF	natural gas	0.00	0	0	0	0
PWP000087	SPWP	natural gas	0.00	0	0	0	0
PWP000088	OGF	natural gas	776.57	0	0	0	0
PWP000089	SPWP	natural gas	4.81	0	0	0	0
PWP000090	Landfill	biomass	148.00	0	0	0	0
PWP000091	SPWP	biomass	0.00	0	0	0	0
PWP000092	SPWP	natural gas	20672.81	0	0	0	0
PWP000093	WWTP	natural gas	127.97	0	0	0	0
PWP000094	SPWP	biomass	0.00	0	0	0	0
PWP000095	SPWP	natural gas	0.00	0	0	0	0
PWP000099	SPWP	natural gas	19400.02	24310	25333.33	0	0
PWP000100	SPWP	biomass	0.00	0	0	0	0
PWP000101	SPWP	biomass	0.00	0	0	0	0
PWP000102	Landfill	biomass	41107.65	840	0	0	0
PWP000103	SPWP	natural gas	145.46	0	0	0	0
PWP000104	SPWP	natural gas	14611.16	12550	13083.33	0	0
PWP000105	SPWP	natural gas	52.91	0	0	0	0
PWP000106	SPWP	natural gas	0.00	0	0	0	0
PWP000107	SPWP	natural gas	22.86	0	0	0	0
PWP000108	SPWP	natural gas	361.37	0	0	0	0
PWP000109	SPWP	natural gas	86.24	100	0	0	0
PWP000110	SPWP	natural gas	691.41	0	0	0	2
PWP000111	OGF	natural gas	1282.97	0	0	0	2
PWP000112	SPWP	natural gas	693.99	0	0	0	2
PWP000114	SPWP	natural gas	81.28	220	0	0	0
PWP000115	WWTP	natural gas	27300.45	3970	4125	0	0
PWP000116	SPWP	biomass	171442.92	0	0	0	0
PWP000117	SPWP	natural gas	0.00	0	0	0	0
PWP000118	SPWP	nuclear	0.00	0	0	0	0
PWP000119	SPWP	biomass	0.00	0	0	0	0
PWP000120	SPWP	natural gas	1196.65	0	0	0	0
PWP000121	Refinery	natural gas	0.00	0	0	5	1
PWP000122	SPWP	natural gas	5935.47	0	0	0	0
PWP000123	SPWP	natural gas	185.97	240	0	0	0

PWP000124	SPWP	natural gas	6.62	0	0	0	0
PWP000125	SPWP	natural gas	0.00	0	0	0	0
PWP000126	SPWP	natural gas	38912.84	17320	0	85	0
PWP000127	SPWP	petroleum	0.00	0	0	0	0
PWP000128	SPWP	natural gas	1521.21	1500	0	0	0
PWP000129	SPWP	natural gas	104.36	150	0	0	0
PWP000130	SPWP	natural gas	5407.43	8820	9166.67	0	0
PWP000131	SPWP	biomass	31489.36	0	0	0	0
PWP000132	Refinery	natural gas	15152.74	11000	11458.33	9	1
PWP000133	SPWP	natural gas	4870.79	0	0	0	1
PWP000134	OGF	natural gas	0.00	0	0	0	0
PWP000135	SPWP	natural gas	28370.85	0	0	20	0
PWP000136	SPWP	natural gas	55.71	0	0	0	0
PWP000138	SPWP	natural gas	0.00	0	0	0	0
PWP000139	WWTP	biomass	4325.38	6320	0	0	0
PWP000140	SPWP	natural gas	68.54	0	0	0	0
PWP000141	Refinery	other	2000.05	0	0	0	1
PWP000142	SPWP	natural gas	64.89	0	0	0	0
PWP000143	SPWP	natural gas	191.22	470	500	0	0
PWP000144	SPWP	natural gas	0.00	0	0	0	0
PWP000145	Refinery	other	0.00	0	0	2	0
PWP000146	Processing Plant	natural gas	0.00	0	0	0	0
PWP000147	SPWP	biomass	45998.42	0	0	0	0
PWP000148	SPWP	natural gas	154.08	0	0	0	0
PWP000149	Refinery	natural gas	7774.81	0	0	0	0
PWP000150	SPWP	natural gas	57.41	0	0	0	0
PWP000151	SPWP	natural gas	68.78	270	0	0	0
PWP000152	SPWP	natural gas	347.91	0	1125	0	0
PWP000153	SPWP	natural gas	909.88	0	0	0	0
PWP000154	Landfill	biomass	141.41	0	0	60	0
PWP000155	Landfill	biomass	308.09	1430	0	0	0
PWP000156	WWTP	biomass	9250.33	0	0	0	0
PWP000157	SPWP	natural gas	20632.77	0	0	0	0
PWP000158	OGF	natural gas	0.00	0	0	0	0
PWP000164	SPWP	natural gas	8.26	0	0	0	0
PWP000165	SPWP	natural gas	333.36	290	375	0	0
PWP000166	SPWP	natural gas	848.24	0	0	0	0
PWP000167	SPWP	natural gas	409.01	860	916.67	0	2
PWP000168	SPWP	natural gas	83.17	320	333.33	0	0
PWP000169	SPWP	natural gas	140.42	0	0	0	0

PWP000170	SPWP	natural gas	101.70	250	0	0	0
PWP000171	SPWP	natural gas	915.19	3490	0	0	0
PWP000172	SPWP	natural gas	77.04	230	250	0	0
PWP000173	SPWP	natural gas	2273.08	2290	2375	0	0
PWP000174	SPWP	biomass	0.00	0	0	0	0
PWP000175	SPWP	natural gas	50.17	0	0	0	0
PWP000176	SPWP	natural gas	137.02	220	250	0	0
PWP000177	SPWP	natural gas	996.04	1000	1083.33	0	0
PWP000178	SPWP	natural gas	0.00	0	0	0	1
PWP000179	SPWP	natural gas	18130.95	21630	26000	7	0
PWP000181	SPWP	natural gas	357.93	430	458.33	0	0
PWP000182	SPWP	petroleum	0.00	0	0	0	0
PWP000183	SPWP	natural gas	24915.31	16540	17250	0	0
PWP000184	SPWP	natural gas	184.36	220	250	0	0
PWP000185	SPWP	natural gas	359.23	0	0	0	0
PWP000186	SPWP	biomass	88654.54	0	0	0	0
PWP000187	SPWP	natural gas	360.86	0	0	0	0
PWP000188	SPWP	natural gas	52.88	0	0	0	0
PWP000189	SPWP	natural gas	461.06	0	791.67	0	0
PWP000190	SPWP	natural gas	3726.27	0	0	0	0
PWP000191	SPWP	natural gas	105.25	740	0	0	0
PWP000192	SPWP	natural gas	0.00	0	0	0	0
PWP000193	SPWP	natural gas	264.67	2680	2791.67	0	0
PWP000194	SPWP	natural gas	64.55	0	0	0	0
PWP000195	Landfill	biomass	0.00	0	0	0	0
PWP000196	SPWP	other	0.00	0	0	0	0
PWP000197	SPWP	natural gas	0.00	0	0	0	0
PWP000199	SPWP	natural gas	61.40	0	0	0	0
PWP000200	SPWP	biomass	44.51	0	0	0	0
PWP000201	SPWP	biomass	60.81	0	0	0	0
PWP000202	SPWP	natural gas	349.60	0	0	0	0
PWP000203	SPWP	natural gas	0.00	0	0	0	0
PWP000204	SPWP	natural gas	61.10	0	0	0	0
PWP000205	SPWP	natural gas	152.57	0	0	0	0
PWP000206	Refinery	natural gas	519.75	0	0	0	0
PWP000207	SPWP	natural gas	8278.89	8690	9041.67	0	1
PWP000208	OGF	natural gas	3240.35	0	0	0	2
PWP000209	Landfill	biomass	906.85	7780	0	0	0
PWP000210	SPWP	natural gas	71.62	0	0	0	0
PWP000211	SPWP	natural gas	1449.85	0	0	0	0

PWP000212	SPWP	petroleum	0.00	0	0	0	0
PWP000213	SPWP	natural gas	32.04	150	166.67	0	0
PWP000214	SPWP	natural gas	155.21	0	0	0	0
PWP000215	SPWP	natural gas	8148.59	21950	22875	14	0
PWP000216	SPWP	natural gas	178.35	0	0	0	0
PWP000217	SPWP	natural gas	153.41	0	0	0	0
PWP000218	SPWP	natural gas	585.50	510	0	0	0
PWP000219	SPWP	natural gas	66.24	0	0	0	0
PWP000220	SPWP	natural gas	62.98	0	0	0	0
PWP000221	Landfill	biomass	313.97	0	0	0	0
PWP000222	Refinery	natural gas	0.00	0	0	0	2
PWP000223	SPWP	natural gas	146.26	230	250	0	0
PWP000224	SPWP	natural gas	56.21	0	0	0	0
PWP000225	SPWP	natural gas	8.74	0	0	0	0
PWP000226	WWTP	natural gas	8921.11	5380	0	0	0
PWP000227	SPWP	natural gas	1035.24	0	1041.67	0	0
PWP000228	SPWP	natural gas	264.54	540	0	0	2
PWP000229	Refinery	other	2145.41	0	0	48	0
PWP000230	SPWP	natural gas	1664.03	1910	2000	15	0
PWP000231	SPWP	natural gas	20797.50	23820	24791.67	0	0
PWP000232	SPWP	natural gas	778.75	0	0	0	2
PWP000233	SPWP	natural gas	12163.67	0	10791.67	0	0
PWP000234	SPWP	natural gas	185.70	220	250	0	0
PWP000235	SPWP	natural gas	5267.29	5610	5833.33	0	0
PWP000238	SPWP	natural gas	0.00	0	0	0	0
PWP000239	Landfill	biomass	0.00	0	0	0	0
PWP000240	SPWP	natural gas	1056.91	1480	1541.67	0	0
PWP000241	SPWP	natural gas	856.24	2030	2125	0	0
PWP000242	Refinery	natural gas	8766.37	7660	8000	0	0
PWP000243	SPWP	other	0.00	0	0	0	0
PWP000244	SPWP	natural gas	64.95	0	0	0	0
PWP000245	SPWP	natural gas	0.00	0	0	0	0
PWP000246	SPWP	natural gas	29.38	0	0	0	0
PWP000247	SPWP	natural gas	23.90	0	0	113	0
PWP000248	SPWP	natural gas	209.75	450	0	0	0
PWP000249	SPWP	natural gas	989.37	0	0	0	0
PWP000250	SPWP	natural gas	0.00	0	0	0	0
PWP000251	SPWP	biomass	0.00	0	0	0	0
PWP000252	SPWP	biomass	0.00	0	0	0	0
PWP000253	SPWP	natural gas	19986.63	15360	16000	0	0

PWP000254	Landfill	biomass	0.00	0	0	0	0
PWP000255	SPWP	natural gas	0.00	0	0	0	0
PWP000256	SPWP	natural gas	615.13	750	791.67	0	0
PWP000257	SPWP	natural gas	10176.30	9960	10375	0	0
PWP000258	SPWP	natural gas	0.00	0	0	0	0
PWP000259	SPWP	other	0.00	0	0	0	0
PWP000260	SPWP	other	0.00	0	0	0	0
PWP000261	SPWP	natural gas	180.93	250	0	0	0
PWP000262	SPWP	natural gas	790.02	1130	1125	0	0
PWP000263	Landfill	biomass	218.20	710	0	0	0
PWP000264	Landfill	biomass	2222.02	960	0	0	0
PWP000265	SPWP	biomass	0.00	0	0	0	0
PWP000266	Landfill	biomass	0.00	0	0	0	0
PWP000267	Landfill	biomass	276.25	280	0	0	0
PWP000268	SPWP	biomass	2435.04	0	0	0	0
PWP000269	SPWP	natural gas	24493.58	30190	31458.33	0	0
PWP000270	SPWP	biomass	0.00	0	0	0	0
PWP000271	SPWP	biomass	145832.95	83340	0	0	0
PWP000272	SPWP	natural gas	46.11	0	0	0	0
PWP000273	SPWP	natural gas	0.00	0	0	0	0
PWP000274	SPWP	natural gas	278.07	0	500	0	0
PWP000275	SPWP	natural gas	0.00	0	0	0	0
PWP000276	SPWP	natural gas	2973.24	0	4458.33	0	0
PWP000277	SPWP	natural gas	848.37	690	708.33	0	0
PWP000279	SPWP	biomass	284.94	0	0	0	0
PWP000280	SPWP	natural gas	0.00	0	0	0	0
PWP000281	OGF	natural gas	0.00	0	0	0	2
PWP000282	SPWP	biomass	0.00	0	0	0	0
PWP000283	SPWP	natural gas	0.00	0	0	0	0
PWP000284	SPWP	natural gas	0.00	0	0	0	0
PWP000285	Refinery	natural gas	0.00	0	0	9	1
PWP000286	Landfill	biomass	0.00	0	0	0	0
PWP000287	SPWP	natural gas	0.00	0	0	0	0
PWP000288	SPWP	natural gas	315.81	0	0	0	0
PWP000289	SPWP	natural gas	36.31	200	2000	0	0
PWP000290	SPWP	natural gas	0.00	0	0	0	0
PWP000291	SPWP	natural gas	224.77	470	500	0	0
PWP000295	SPWP	natural gas	2090.14	2860	3000	0	0
PWP000296	SPWP	natural gas	2.29	0	0	0	0
PWP000297	SPWP	other	0.00	0	0	0	0

PWP000298	Landfill	biomass	373.74	1560	0	0	0
PWP000299	SPWP	natural gas	5497.42	15610	0	0	0
PWP000300	SPWP	natural gas	5591.83	0	0	0	0
PWP000301	SPWP	natural gas	2164.74	0	0	0	0
PWP000302	WWTP	biomass	1975.67	0	0	0	0
PWP000303	SPWP	natural gas	403.62	0	0	0	0
PWP000304	SPWP	biomass	0.00	0	0	0	0
PWP000305	Digester	biomass	0.00	0	0	58	0
PWP000306	SPWP	biomass	68112.96	45290	0	0	0
PWP000307	SPWP	other	0.00	0	0	0	0
PWP000308	SPWP	natural gas	8817.72	18190	0	0	0
PWP000309	SPWP	natural gas	6247.76	0	0	0	0
PWP000310	SPWP	natural gas	38.56	0	0	0	0
PWP000311	Refinery	natural gas	0.00	0	0	0	0
PWP000312	SPWP	natural gas	942.39	0	0	0	0
PWP000313	SPWP	natural gas	30190.18	27570	28708.33	0	0
PWP000314	SPWP	natural gas	1952.97	0	0	0	0
PWP000315	SPWP	petroleum	0.00	0	0	0	0
PWP000316	SPWP	petroleum	0.00	0	0	0	0
PWP000317	Refinery	other	2137.61	0	0	0	0
PWP000318	SPWP	natural gas	1122.87	1700	1750	0	0
PWP000319	SPWP	natural gas	0.00	0	0	0	0
PWP000320	SPWP	natural gas	0.00	0	0	0	0
PWP000321	SPWP	natural gas	64.50	0	0	0	0
PWP000322	WWTP	biomass	0.00	0	0	0	0
PWP000323	WWTP	biomass	0.00	0	0	2	0
PWP000325	Landfill	biomass	688.59	2170	0	0	0
PWP000326	WWTP	biomass	9498.04	0	0	0	0
PWP000327	SPWP	petroleum	0.00	0	0	0	0
PWP000328	Landfill	biomass	538.27	0	0	0	0
PWP000329	SPWP	biomass	2534.67	9040	0	0	0
PWP000330	SPWP	biomass	0.00	0	0	0	0
PWP000331	SPWP	natural gas	379.91	0	0	0	0
PWP000332	Landfill	biomass	0.00	0	0	0	0
PWP000333	SPWP	natural gas	130.48	0	0	0	0
PWP000334	SPWP	natural gas	2045.53	1660	1750	0	0
PWP000335	WWTP	natural gas	4959.96	0	0	0	0
PWP000336	SPWP	natural gas	110.81	0	0	0	0
PWP000337	Refinery	natural gas	10038.46	0	0	0	0
PWP000338	SPWP	biomass	86804.62	59190	0	0	0

PWP000339	SPWP	coal	0.00	0	0	0	0
PWP000340	SPWP	biomass	81905.14	68000	0	0	0
PWP000341	SPWP	natural gas	160.24	930	958.33	0	0
PWP000342	WWTP	natural gas	886.69	1010	1041.67	0	0
PWP000343	WWTP	biomass	3104.25	0	0	0	0
PWP000344	SPWP	natural gas	234.29	0	0	0	0
PWP000345	SPWP	natural gas	9.75	0	0	0	0
PWP000346	SPWP	natural gas	0.00	0	0	0	0
PWP000349	SPWP	natural gas	48.04	0	0	0	0
PWP000350	SPWP	biomass	0.00	0	0	0	0
PWP000351	SPWP	natural gas	28.94	0	0	0	0
PWP000352	SPWP	natural gas	2424.20	290	0	0	0
PWP000353	WWTP	biomass	0.00	0	0	0	0
PWP000354	WWTP	natural gas	7000.60	5780	6000	0	0
PWP000355	SPWP	natural gas	453.95	0	0	0	0
PWP000356	SPWP	natural gas	0.00	0	0	0	0
PWP000357	OGF	natural gas	4130.17	1240	0	0	0
PWP000363	SPWP	natural gas	63.51	0	0	0	0
PWP000364	SPWP	natural gas	712.76	0	791.67	0	0
PWP000365	SPWP	natural gas	556.92	0	0	0	0
PWP000366	SPWP	natural gas	792.96	0	0	0	0
PWP000367	SPWP	biomass	123.92	400	0	0	0
PWP000368	Refinery	other	0.00	0	0	0	0
PWP000369	SPWP	biomass	0.00	0	0	0	0
PWP000370	Landfill	other	85.75	0	0	0	0
PWP000371	Landfill	natural gas	92.32	0	0	0	0
PWP000372	SPWP	natural gas	0.00	0	0	0	0
PWP000373	SPWP	petroleum	0.00	0	0	0	0
PWP000374	SPWP	natural gas	6392.04	5700	5916.67	0	0
PWP000375	WWTP	natural gas	7581.77	15410	16083.33	0	2
PWP000376	SPWP	biomass	0.00	0	0	0	0
PWP000378	SPWP	natural gas	5984.40	4810	0	0	0
PWP000379	SPWP	biomass	1220.30	0	0	0	0
PWP000380	SPWP	natural gas	90.92	0	0	0	0
PWP000381	SPWP	biomass	0.00	0	0	0	0
PWP000382	SPWP	natural gas	103.53	0	0	0	0
PWP000383	SPWP	biomass	0.00	0	0	0	0
PWP000384	SPWP	biomass	54954.68	0	0	0	0
PWP000385	SPWP	biomass	77237.48	0	0	0	0
PWP000386	SPWP	biomass	119414.20	0	0	0	0

PWP000387	SPWP	biomass	42427.40	0	0	0	0
PWP000388	Processing Plant	other	0.00	0	0	21	0
PWP000389	WWTP	biomass	2358.19	750	0	0	0
PWP000390	WWTP	natural gas	11035.81	0	0	0	0
PWP000391	SPWP	natural gas	0.00	0	0	0	0
PWP000393	Landfill	biomass	206.08	0	0	0	0
PWP000394	Landfill	biomass	191.44	0	0	0	0
PWP000395	Landfill	biomass	12.94	0	0	0	0
PWP000396	SPWP	natural gas	27.03	0	0	0	0
PWP000397	SPWP	biomass	2100.92	0	0	0	0
PWP000398	SPWP	natural gas	5360.67	0	0	0	2
PWP000399	OGF	natural gas	1396.60	0	0	0	2
PWP000400	SPWP	biomass	48332.24	0	0	0	0
PWP000401	SPWP	natural gas	5509.26	0	0	0	0
PWP000402	Landfill	biomass	0.00	0	0	0	0
PWP000403	SPWP	biomass	98438.36	0	0	0	0
PWP000404	SPWP	natural gas	1286.55	0	0	0	0
PWP000405	SPWP	natural gas	11.60	0	0	0	0
PWP000406	SPWP	natural gas	503.40	460	0	0	0
PWP000407	SPWP	natural gas	61.06	0	0	0	0
PWP000408	SPWP	biomass	187108.22	0	0	0	0
PWP000409	WWTP	biomass	0.00	0	0	0	0
PWP000410	WWTP	biomass	1203.34	0	0	0	0
PWP000411	SPWP	natural gas	15060.77	20350	21208	0	0
PWP000412	Landfill	biomass	1928.93	5780	0	0	0
PWP000413	SPWP	natural gas	5263.41	0	0	0	0
PWP000414	SPWP	natural gas	7436.58	0	0	0	1
PWP000415	SPWP	biomass	454.81	0	0	0	0
PWP000416	OGF	natural gas	1289.64	0	0	0	0
PWP000417	Landfill	biomass	288.13	0	0	0	0
PWP000418	SPWP	natural gas	32.86	0	0	0	0
PWP000419	SPWP	other	0.00	0	0	0	0
PWP000420	Refinery	other	0.00	0	0	2	1
PWP000421	SPWP	biomass	0.00	0	0	0	0
PWP000422	OGF	natural gas	0.00	0	0	0	0
PWP000423	SPWP	natural gas	3401.88	0	0	0	2
PWP000424	WWTP	biomass	51492.35	0	0	18	0
PWP000425	Landfill	biomass	0.00	0	0	0	0
PWP000426	SPWP	natural gas	7142.18	7100	7416.67	0	0
PWP000427	SPWP	natural gas	346.09	0	0	0	0

PWP000428	SPWP	natural gas	1280.76	0	0	0	0
PWP000429	WWTP	natural gas	75.53	0	0	0	0
PWP000430	SPWP	natural gas	2570.02	0	4500	0	0
PWP000431	SPWP	biomass	4936.92	0	0	0	0
PWP000432	SPWP	natural gas	363.74	0	0	0	0
PWP000433	SPWP	natural gas	125.87	0	0	0	0
PWP000434	SPWP	natural gas	2688.59	0	0	0	0
PWP000435	Landfill	natural gas	4235.74	0	0	0	0
PWP000436	WWTP	biomass	0.00	0	0	0	0
PWP000437	SPWP	other	0.00	0	0	0	0
PWP000438	Refinery	natural gas	3107.75	0	0	0	0
PWP000439	SPWP	natural gas	17182.13	11140	12500	0	0
PWP000440	SPWP	natural gas	60.99	0	0	0	0
PWP000441	SPWP	petroleum	0.00	0	0	0	0
PWP000442	Digester	biomass	0.00	0	0	0	0
PWP000443	SPWP	petroleum	0.00	0	0	0	0
PWP000445	SPWP	natural gas	378.61	0	0	0	0
PWP000446	SPWP	biomass	10.39	0	0	0	0
PWP000447	SPWP	natural gas	4.68	0	0	0	0
PWP000448	SPWP	natural gas	4717.60	3390	3416.67	0	0
PWP000449	SPWP	natural gas	8578.38	9340	9750	0	0
PWP000450	SPWP	natural gas	52.47	0	0	0	0
PWP000451	Refinery	natural gas	29648.63	0	0	0	0
PWP000452	SPWP	natural gas	608.11	0	0	0	2
PWP000453	SPWP	natural gas	0.00	0	0	0	0
PWP000454	SPWP	natural gas	1973.07	0	0	0	1
PWP000455	SPWP	natural gas	0.00	0	0	0	2
PWP000456	SPWP	biomass	162871.08	25160	0	0	0
PWP000457	SPWP	biomass	0.00	0	0	0	0
PWP000458	Refinery	other	0.00	0	0	0	2
PWP000459	SPWP	natural gas	88.46	0	0	0	0
PWP000460	SPWP	natural gas	3148.52	0	0	0	0
PWP000461	SPWP	biomass	98929.85	0	0	0	0
PWP000462	WWTP	biomass	8089.63	0	0	0	0
PWP000463	SPWP	natural gas	67.18	0	0	0	0
PWP000464	SPWP	natural gas	61.13	0	0	0	0
PWP000465	SPWP	other	0.00	0	0	0	0
PWP000466	SPWP	natural gas	355.93	280	0	0	0
PWP000467	SPWP	natural gas	344.88	460	500	0	0
PWP000468	Landfill	biomass	0.00	0	0	0	0

Table S4.4 HyTES Power Plant GSAAM attributed raw data. This table provides the emissions for plumes that were observed by HyTES that were attributed to each power plant in the Vista-CA dataset.

Facility Vista-ID	Sub-Facility Vista-ID	Year	Month	Day	Time	Area	Latitude	Longitude
PWP000029		2014	7	8	165610	Trona	35.765495	-117.38161
PWP000029		2014	7	8	165610	Trona	35.765466	-117.38433
PWP000029		2014	7	8	165610	Trona	35.76546	-117.38167
PWP000029		2014	7	8	165610	Trona	35.765408	-117.38427
PWP000029		2014	7	8	165610	Trona	35.765318	-117.38421
PWP000029		2014	7	8	165610	Trona	35.765392	-117.38454
PWP000029		2014	7	8	165610	Trona	35.765338	-117.38406
PWP000029		2014	7	8	165610	Trona	35.765297	-117.38402
PWP000029		2014	7	8	165610	Trona	35.765263	-117.38398
PWP000029		2014	7	8	165610	Trona	35.765246	-117.38336
PWP000029		2014	7	8	165610	Trona	35.765234	-117.38402
PWP000029		2014	7	8	165610	Trona	35.765223	-117.3841
PWP000029		2014	7	8	165610	Trona	35.765164	-117.38398
PWP000029		2014	7	8	165610	Trona	35.765164	-117.38331
PWP000029		2014	7	8	165610	Trona	35.765142	-117.38318
PWP000029		2014	7	8	165610	Trona	35.76506	-117.38368
PWP000029		2014	7	8	165610	Trona	35.765042	-117.38322
PWP000029		2014	7	8	165610	Trona	35.765038	-117.3831
PWP000029		2014	7	8	165610	Trona	35.765018	-117.38375
PWP000029		2014	7	8	165610	Trona	35.76499	-117.3838
PWP000029		2014	7	8	165610	Trona	35.764955	-117.38342
PWP000029		2014	7	8	165610	Trona	35.764955	-117.38337
PWP000029		2014	7	8	165610	Trona	35.764076	-117.38371
PWP000029		2014	7	8	165610	Trona	35.764039	-117.38374
PWP000029		2014	7	8	165610	Trona	35.763998	-117.38378
PWP000029		2014	7	8	165610	Trona	35.763981	-117.38331
PWP000029		2014	7	8	165610	Trona	35.764025	-117.38571
PWP000029		2014	7	8	165610	Trona	35.763949	-117.38335
PWP000029		2014	7	8	165610	Trona	35.763902	-117.38333
PWP000029		2014	7	8	165610	Trona	35.763746	-117.38496
PWP000029		2014	7	8	165610	Trona	35.76372	-117.38498
PWP000029		2014	7	8	165610	Trona	35.763702	-117.38502
PWP000029		2014	7	8	165610	Trona	35.763702	-117.38397
PWP000029		2014	7	8	165610	Trona	35.763654	-117.3839
PWP000029		2014	7	8	165610	Trona	35.76278	-117.38401

PWP000033		2015	2	5	202724	KernRiverOil	35.418996	-118.92632
PWP000033		2015	2	8	234132	KernRiverOil	35.418918	-118.9263
PWP000033		2015	2	9	193430	KernRiverOil	35.418937	-118.9263
PWP000046		2014	7	5	193259	LongBeachNG	33.82409	-118.24108
PWP000046		2014	7	5	193259	LongBeachNG	33.823808	-118.23866
PWP000046		2014	7	5	193259	LongBeachNG	33.823307	-118.23982
PWP000046		2014	7	5	193259	LongBeachNG	33.823024	-118.23948
PWP000046		2014	7	5	193259	LongBeachNG	33.822949	-118.23964
PWP000046		2014	7	5	193259	LongBeachNG	33.822953	-118.23948
PWP000046		2014	7	5	193259	LongBeachNG	33.822543	-118.24091
PWP000046		2014	7	5	193259	LongBeachNG	33.822454	-118.24084
PWP000046		2014	7	5	193259	LongBeachNG	33.822418	-118.24076
PWP000046		2014	7	5	193259	LongBeachNG	33.822129	-118.23943
PWP000046		2014	7	5	193259	LongBeachNG	33.822062	-118.23977
PWP000046		2014	7	5	193259	LongBeachNG	33.821813	-118.23952
PWP000046		2014	7	5	193259	LongBeachNG	33.816279	-118.23306
PWP000121		2014	7	5	192352	LongBeachNG	33.845847	-118.23337
PWP000121		2014	7	5	192352	LongBeachNG	33.845606	-118.23336
PWP000121		2014	7	5	192352	LongBeachNG	33.845259	-118.2333
PWP000121		2014	7	5	192352	LongBeachNG	33.843416	-118.23418
PWP000121		2014	7	5	192352	LongBeachNG	33.843382	-118.233
PWP000121		2014	7	5	192352	LongBeachNG	33.843246	-118.23314
PWP000121		2014	7	5	192352	LongBeachNG	33.843165	-118.23343
PWP000121		2014	7	5	192352	LongBeachNG	33.843164	-118.23368
PWP000121		2014	7	5	192352	LongBeachNG	33.8428	-118.23218
PWP000121		2014	7	5	192352	LongBeachNG	33.841767	-118.23535
PWP000121		2014	7	5	192352	LongBeachNG	33.841387	-118.23568
PWP000133		2014	7	5	194225	ElSegundo	33.910636	-118.42559
PWP000133		2014	7	5	194225	ElSegundo	33.910652	-118.42544
PWP000133		2014	7	5	194225	ElSegundo	33.910396	-118.42421
PWP000133		2014	7	5	194225	ElSegundo	33.910369	-118.42417
PWP000133		2014	7	5	194225	ElSegundo	33.910425	-118.42556
PWP000133		2014	7	5	194225	ElSegundo	33.910401	-118.42405
PWP000133		2014	7	5	194225	ElSegundo	33.910464	-118.4251
PWP000133		2014	7	5	194225	ElSegundo	33.910239	-118.42474
PWP000133		2014	7	5	194225	ElSegundo	33.910173	-118.4245
PWP000133		2014	7	5	194225	ElSegundo	33.910084	-118.42448
PWP000133		2014	7	5	194225	ElSegundo	33.910122	-118.4239
PWP000133		2014	7	5	194225	ElSegundo	33.909549	-118.42395
PWP000133		2014	7	5	194225	ElSegundo	33.909731	-118.42472

PWP000133		2014	7	5	194225	ElSegundo	33.909507	-118.42403
PWP000133		2014	7	5	194225	ElSegundo	33.909443	-118.4242
PWP000133		2014	7	5	194225	ElSegundo	33.909399	-118.42404
PWP000133		2014	7	5	194225	ElSegundo	33.90931	-118.42403
PWP000133		2014	7	5	194225	ElSegundo	33.909628	-118.42502
PWP000133		2014	7	5	194225	ElSegundo	33.909265	-118.42382
PWP000133		2014	7	5	194225	ElSegundo	33.909219	-118.42419
PWP000133		2014	7	5	194225	ElSegundo	33.909236	-118.42451
PWP000133		2014	7	5	195005	ElSegundo	33.911085	-118.42567
PWP000133		2014	7	5	195005	ElSegundo	33.910489	-118.42471
PWP000133		2014	7	5	195005	ElSegundo	33.910434	-118.42526
PWP000133		2014	7	5	195005	ElSegundo	33.909997	-118.42437
PWP000133		2014	7	5	195005	ElSegundo	33.909281	-118.42413
PWP000133		2014	7	5	195005	ElSegundo	33.909278	-118.42435
PWP000133		2014	7	5	195005	ElSegundo	33.909184	-118.42467
PWP000133		2014	7	5	195005	ElSegundo	33.909012	-118.42401
PWP000133		2014	7	5	195005	ElSegundo	33.909016	-118.42394
PWP000133		2016	1	25	202848	Hyperion	33.911758	-118.42568
PWP000133		2016	1	25	202848	Hyperion	33.911246	-118.42574
PWP000133		2016	1	25	202848	Hyperion	33.911126	-118.4259
PWP000133		2016	1	25	202848	Hyperion	33.910708	-118.42558
PWP000133		2016	1	25	202848	Hyperion	33.910368	-118.42404
PWP000207		2015	2	8	231927	KernRiverOil	35.453538	-118.98505
PWP000207		2015	2	8	231927	KernRiverOil	35.453551	-118.98467
PWP000285		2015	2	8	222449	KernRiverOil	35.419493	-119.01029
PWP000285		2015	2	8	222449	KernRiverOil	35.419456	-119.01189
PWP000285		2015	2	8	222449	KernRiverOil	35.419344	-119.01104
PWP000414		2014	7	8	191611	KernRiverOil	35.454987	-119.00159
PWP000414		2014	7	8	191611	KernRiverOil	35.454918	-119.00167
PWP000414		2015	2	8	225750	KernRiverOil	35.454887	-119.00167
PWP000414		2015	2	8	225750	KernRiverOil	35.454549	-119.00096
PWP000414		2015	2	8	225750	KernRiverOil	35.454506	-119.00169
PWP000414		2015	2	8	225750	KernRiverOil	35.454453	-119.00364
PWP000414		2015	2	9	185105	KernRiverOil	35.454784	-119.00197
PWP000454		2015	2	8	223542	KernRiverOil	35.440369	-119.01285
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906979	-118.40322
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906856	-118.40402
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906852	-118.40312
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906759	-118.40357
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906538	-118.40369

REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906574	-118.40407
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906437	-118.40542
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906444	-118.40316
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906238	-118.40331
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906188	-118.40527
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906193	-118.40329
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906146	-118.40475
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.90605	-118.40653
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906077	-118.40607
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906024	-118.40621
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906197	-118.40448
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.906009	-118.40507
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905931	-118.40496
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905942	-118.4028
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905819	-118.4022
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905801	-118.4021
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905703	-118.40294
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905689	-118.40707
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905703	-118.4045
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905719	-118.40305
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905668	-118.40519
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905679	-118.40669
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.90565	-118.40703
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905645	-118.40296
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905617	-118.40439
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905618	-118.40479
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905602	-118.4039
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905534	-118.40386
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905506	-118.40363
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.905668	-118.40569
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.90528	-118.40346
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904679	-118.40259
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904717	-118.40606
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904606	-118.40363
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904582	-118.40262
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904382	-118.40348
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904308	-118.40443
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.90435	-118.40351
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904225	-118.4044
REF000005	PWP000132	2014	7	5	194225	ElSegundo	33.904146	-118.40418

REF000018	PWP000121	2014	7	5	192352	LongBeachNG	33.843834	-118.23635
REF000018	PWP000121	2014	7	5	192352	LongBeachNG	33.842466	-118.23693
REF000018	PWP000121	2014	7	5	192352	LongBeachNG	33.842301	-118.23695
REF000020	PWP000141	2014	7	5	193259	LongBeachNG	33.791326	-118.2325
REF000020	PWP000141	2014	7	5	193259	LongBeachNG	33.791185	-118.23374
REF000020	PWP000141	2014	7	5	193259	LongBeachNG	33.790135	-118.23412
REF000020	PWP000141	2014	7	5	193259	LongBeachNG	33.790135	-118.23374
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.820655	-118.23982
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.820343	-118.23968
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.81961	-118.23982
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.819232	-118.23997
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.819039	-118.23514
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.818745	-118.23532
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.818463	-118.23663
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.818	-118.23875
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.817997	-118.23539
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.817893	-118.23569
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.817655	-118.23543
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.817589	-118.23581
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.817455	-118.23535
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.817451	-118.23559
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.816748	-118.2355
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.816413	-118.23381
REF000021	PWP000046	2014	7	5	193259	LongBeachNG	33.816396	-118.23642
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.419166	-119.01179
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.419033	-119.01179
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.419061	-119.01167
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418984	-119.01129
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418903	-119.01172
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418892	-119.01164
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418483	-119.01106
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.41844	-119.01115
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418325	-119.00902
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418299	-119.00904
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418296	-119.01201
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418273	-119.00951
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418267	-119.00901
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418255	-119.00896
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.41815	-119.01197
REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.418187	-119.00958

REF000023	PWP000285	2015	2	5	190531	KernRiverOil	35.416974	-119.00959
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.419218	-119.01149
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.419175	-119.01144
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.419087	-119.01144
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.419061	-119.01147
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418967	-119.01147
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418885	-119.0111
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418875	-119.01126
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418845	-119.01122
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418816	-119.01109
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418748	-119.01116
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.41871	-119.0114
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418719	-119.01132
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418608	-119.01135
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.41836	-119.01118
REF000023	PWP000285	2015	2	8	221409	KernRiverOil	35.418343	-119.01053
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.419043	-119.00829
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.419008	-119.01099
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.419136	-119.01121
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418921	-119.01118
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.41892	-119.01074
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418886	-119.01023
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418772	-119.01099
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418731	-119.00884
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418599	-119.00983
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418528	-119.0104
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.4187	-119.01109
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418274	-119.01017
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418267	-119.01093
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418115	-119.00892
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418116	-119.00817
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.418091	-119.00883
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417835	-119.00903
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417685	-119.01126
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417494	-119.01177
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417459	-119.01179
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417483	-119.00896
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417423	-119.00932
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417332	-119.00923
REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.417344	-119.01147

REF000023	PWP000285	2015	2	8	222449	KernRiverOil	35.416728	-119.01174
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.419143	-119.01068
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.419061	-119.01106
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.419038	-119.01079
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.419061	-119.01073
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.418912	-119.01121
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.418746	-119.01093
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.418644	-119.01103
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.418607	-119.01096
REF000023	PWP000285	2015	2	9	181750	KernRiverOil	35.418401	-119.01095
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.778626	-118.22838
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.778523	-118.22874
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.778299	-118.22869
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.778293	-118.22835
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.77794	-118.22903
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.777345	-118.22782
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.777236	-118.22776
REF000025	PWP000178	2014	7	5	193259	LongBeachNG	33.777037	-118.22991
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.776951	-118.22876
REF000025	PWP000178	2014	7	5	193259	LongBeachNG	33.776967	-118.22975
REF000025	PWP000178	2014	7	5	193259	LongBeachNG	33.776452	-118.23054
REF000025	PWP000178	2014	7	5	193259	LongBeachNG	33.776417	-118.23061
REF000025	PWP000178	2014	7	5	193259	LongBeachNG	33.776381	-118.23072
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.776265	-118.22862
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.775979	-118.22809
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.776067	-118.22858
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.775953	-118.2287
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.775708	-118.22807
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.775704	-118.22827
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.77549	-118.22833
REF000025	PWP000420	2014	7	5	193259	LongBeachNG	33.775455	-118.22838

Table S4.5. AVIRIS-NG Power Plant GSAAM attributed emissions raw data. This table provides the emissions for each plume observed by AVIRIS-NG along with where and when it was captured for each power plant in the Vista-CA dataset.

Facility Vista-ID	Sub-Facility Vista-ID	Plume-ID	Latitude	Longitude	Year	Month	Day	Time	Emission kg CH_4 h^{-1}
DAF001287	PWP000305	_000136	36.032425	-119.510296	2016	9	13	213245	234
DAF001287	PWP000305	_000137	36.032425	-119.510296	2016	10	26	173255	0
DAF001287	PWP000305	_000138	36.032425	-119.510296	2016	10	26	180214	0
DAF001287	PWP000305	_000443	36.032425	-119.510296	2017	6	16	212046	272
DAF001287	PWP000305	_000444	36.032425	-119.510296	2017	9	20	193250	183
DAF001287	PWP000305	_000445	36.032425	-119.510296	2017	9	20	194154	235
DAF001287	PWP000305	_000446	36.032425	-119.510296	2017	9	28	195131	0
FAB002088	PWP000037	_000065	34.385725	-118.496705	2016	9	15	180022	78
FAB002088	PWP000037	_000083	34.385157	-118.4963171	2016	10	3	193609	136
FAB002249	PWP000135	_000544	35.280734	-119.4726061	2017	9	6	210217	868
LNF000022	PWP000015	_001094	37.749203	-121.650146	2017	6	18	183142	45
LNF000516	PWP000154	_000584	38.314634	-121.8337761	2017	6	17	185116	962
LNF000714	PWP000021	_001125	36.53149	-121.406066	2017	10	7	204208	52
NPP000023	PWP000388	_000898	33.811204	-118.174816	2017	9	1	221701	528
PWP000006		_000887	33.851096	-118.394712	2017	9	1	214553	81
PWP000051		_000895	32.694343	-117.146467	2017	9	25	212203	515
PWP000126		_000877	36.804618	-121.777183	2017	10	7	201538	85
PWP000135		_000228	35.280645	-119.4783614	2016	10	29	205722	146
PWP000135		_000285	35.281571	-119.4808156	2016	10	29	205722	253
PWP000135		_000478	35.280645	-119.4783614	2017	6	15	233419	0
PWP000135		_000479	35.280645	-119.4783614	2017	6	20	164237	27
PWP000135		_000480	35.280645	-119.4783614	2017	6	20	164926	24
PWP000135		_000481	35.280645	-119.4783614	2017	6	20	165600	0
PWP000135		_000503	35.281571	-119.4808156	2017	6	15	230609	199
PWP000135		_000504	35.281571	-119.4808156	2017	6	20	164237	43
PWP000135		_000505	35.281571	-119.4808156	2017	6	20	164926	65
PWP000135		_000506	35.281571	-119.4808156	2017	6	20	165600	47
PWP000135		_000545	35.279019	-119.4760295	2017	6	20	164237	50
PWP000135		_000546	35.279019	-119.4760295	2017	9	6	210217	476
PWP000135		_001246	35.279259	-119.476386	2018	9	19	204226	0
PWP000135		_001247	35.279259	-119.476386	2018	9	19	204827	0
PWP000179		_000595	33.765019	-118.0992873	2017	8	30	191500	75
PWP000179		_000596	33.765019	-118.0992873	2017	9	1	221701	105
PWP000215		_000699	35.295576	-119.5921516	2017	9	26	224352	515

PWP000230		_000714	37.424368	-121.932121	2017	10	5	191421	50
PWP000230		_000715	37.424368	-121.932121	2017	10	5	194644	194
PWP000230		_000716	37.424368	-121.932121	2017	10	5	195334	86
PWP000247		_000876	37.627517	-120.93185	2017	9	30	214843	113
REF000005	PWP000132	_000018	33.905431	-118.406435	2016	9	10	193531	49
REF000005	PWP000132	_000019	33.905983	-118.403489	2016	9	10	193531	0
REF000005	PWP000132	_000368	33.906	-118.4063	2017	3	9	221021	248
REF000005	PWP000132	_000369	33.906	-118.4063	2017	6	16	183155	88
REF000005	PWP000132	_000370	33.905431	-118.406435	2017	9	1	212236	166
REF000005	PWP000132	_000792	33.905751	-118.406747	2017	10	24	182513	0
REF000006	PWP000145	_000879	33.851301	-118.331618	2017	10	24	184733	36
REF000006	PWP000145	_000880	33.851301	-118.331618	2017	10	24	210834	76
REF000006	PWP000145	_000881	33.851301	-118.331618	2017	10	24	213114	32
REF000013	PWP000229	_000020	33.776957	-118.28704	2016	9	10	201644	54
REF000013	PWP000229	_000371	33.772601	-118.286431	2017	9	1	211100	77
REF000013	PWP000229	_000608	33.7751	-118.288828	2017	10	23	202450	314
REF000013	PWP000229	_000609	33.7751	-118.288828	2017	10	24	203615	180
REF000013	PWP000229	_000610	33.774404	-118.282011	2017	9	1	211100	37
REF000013	PWP000229	_000919	33.774727	-118.290496	2017	10	24	183630	783
REF000013	PWP000229	_001182	33.772601	-118.286431	2018	10	1	183753	0
REF000013	PWP000229	_001183	33.772601	-118.286431	2018	10	1	184401	0
REF000013	PWP000229	_001274	33.775293	-118.288664	2018	9	16	204323	0
REF000013	PWP000229	_001275	33.775224	-118.288657	2018	9	16	212943	0
REF000013	PWP000229	_001596	33.775195	-118.289656	2018	10	1	183753	0
REF000013	PWP000229	_001597	33.775195	-118.289656	2018	10	1	184401	0
REF000013	PWP000229	_001598	33.775385	-118.288712	2018	10	1	183753	0
REF000013	PWP000229	_001599	33.775385	-118.288712	2018	10	1	184401	0
REF000018	PWP000121	_000778	33.842039	-118.237465	2017	9	1	201246	52
REF000018	PWP000121	_000839	33.845592	-118.238142	2017	9	1	221701	856
REF000019	PWP000242	_001083	38.014542	-122.11423	2017	9	10	200709	0
REF000021	PWP000046	_000122	33.816222	-118.235695	2016	9	10	205021	0
REF000021	PWP000046	_000432	33.816222	-118.235695	2017	3	9	221021	220
REF000021	PWP000046	_000781	33.817154	-118.235458	2017	9	1	202358	84
REF000021	PWP000046	_000782	33.817154	-118.235458	2017	10	23	210905	0
REF000021	PWP000046	_000783	33.817154	-118.235458	2017	10	24	192146	162
REF000021	PWP000046	_000945	33.818714	-118.236133	2017	3	9	221021	257
REF000021	PWP000046	_001360	33.81682	-118.23507	2018	9	16	211358	0
REF000021	PWP000046	_001361	33.816777	-118.234973	2018	9	16	215953	0
REF000023	PWP000285	_000679	35.419299	-119.0083383	2017	3	9	195546	135
REF000023	PWP000285	_000680	35.419299	-119.0083383	2017	9	8	205824	223

REF000023	PWP000285	_000681	35.419299	-119.0083383	2017	9	26	205517	123
REF000025	PWP000420	_000925	33.776328	-118.22919	2017	3	9	222247	173
REF000025	PWP000420	_000926	33.776328	-118.22919	2017	10	23	205826	51
WWT000059	PWP000424	_000930	33.803642	-118.284674	2017	9	1	205911	480
WWT000059	PWP000424	_000931	33.803642	-118.284674	2017	10	24	204657	358
WWT000059	PWP000424	_000932	33.803642	-118.284674	2017	10	24	215215	148
WWT000087	PWP000323	_000597	33.639371	-117.9565514	2017	8	30	192655	115
WWT000087	PWP000323	_000805	33.638199	-117.956988	2017	9	22	224720	0

VI. References

1. Intergovernmental Panel on Climate Change. 2014. Climate Change 2014: Synthesis Report. https://archive.ipcc.ch/report/ar5/syr/. (accessed 13th, October 2021).

2. Energy Information Administration Independent Statistics & Analysis. Layer Information for Interactive State Maps. 2019. https://www.eia.gov/maps/layer_info-m.php. (accessed 13th, October 2021).

3. California Energy Commission. 2018 Total System Electric Generation. 2019. https://www.energy.ca.gov/data-reports/energy-almanac/california-electricity-data/2019-total-system-electric-generation/2018#:~:text=California%20has%20approximately%2080%2C000%20MWand%206%2C000%20MW%20from%20wind. (accessed 13th, October 2021).

4. Alvarez, R. A., Pacala, S. W., Winebrake, J. J., Chameides, W. L., Hamburg. S. P. 2012. Methane leakage from natural gas infrastructure. Proceedings of the National Academy of Sciences, 109, 6435-6440, DOI: http://doi.org/10.1073/pnas.1202407109.

5. Jeong, S., Millstein, D., and Fischer, M. L. 2014. Spatially Explicit Methane Emissions from Petroleum Production and the Natural Gas System in California. Environmental Science & Technology, 48, 5982–5990, DOI: https://doi.org/10.1021/es4046692.

6. Lara. SB-1383 Short-lived climate pollutants: methane emissions: dairy and livestock: organic waste: landfills. 2015. California Legislative Information. https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB1383. (accessed 13th, October 2021).

7. Nunez and Pavley. AB-32 Air pollution: greenhouse gases: California Global Warming Solutions Act of 2006. California Legislative Information. 2006. https://leginfo.legislature .ca.gov/faces/billNavClient.xhtml?bill_id=200520060AB32. (accessed 13th, October 2021).

8. Thurmond. AB-1496 Methane emissions (2015-2016). 2015. California Legislative Information. https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id= 201520160AB1496. (accessed 13th, October 2021).

9. Ayasse, A.K., Dennison, P.E., Foote, M., Thorpe, A.K., Joshi, S., Green, R.O., Duren, R.M., Thompson, D.R. and Roberts, D.A. 2019. Methane Mapping with Future Satellite Imaging Spectrometers. Remote Sensing, 11(24), p.3054, DOI: https://doi.org/10.3390/rs11243054

10. Crisp, D., Meijer, Y., Munro, R., Bowman, K., Chatterjee, A., Baker, D., Chevallier F., Nassar R., Palmer, P. I., Agusti-Panareda, A., Al-Saadi, J., Ariel, Y., Basu, S., Bergamaschi, P., Boesch, H., Bousquet, P., Bovensmann, H., Bréon, F.-M., Brunner, D., Buchwitz, M., Buisson, F., Burrows, J. P., Butz, A., Ciais, P., Clerbaux, C., Counet, P., Crevoisier, C., Crowell, S., DeCola, P. L., Deniel, C., Dowell, M., Eckman, R., Edwards, D., Ehret, G., Eldering, A., Engelen, R., Fisher, B., Germain, S., Hakkarainen, J., Hilsenrath, E., Holmlund, K., Houweling, S., Hu, H., Jacob, D., Janssens-Maenhout, G., Jones, D., Jouglet, D., Kataoka, F., Kiel, M., Kulawik, S. S., Kuze, S., Lachance, R. L., Lang, R., Landgraf, J., Liu, J., Liu, Y., Maksyutov, S., Matsunaga, T., McKeever, J., Moore, B., Nakajima, M., Natraj, V., Nelson, R. R., Niwa, Y., Oda, T., O'Dell, C. W., Ott, L., Patra, P., Pawson, S., Payne, V., Pinty, B., Polavarapu, S. M., Retscher, C., Rosenberg, R., Schuh, A., Schwandner, F. M., Shiomi, K., Su, W., Tamminen, J., Taylor, T. E., Veefkind, P., Veihelmann, B., Wofsy, S., Worden, J., Wunch, D., Yang, D., Zhang, P., Zehner, C. 2018. A constellation architecture for monitoring carbon dioxide and methane from space. Prepared by the CEOS Atmospheric Constellation Greenhouse Gas Team, Version, 1(8).

11. Cusworth, D.H., Jacob, D.J., Varon, D.J., Miller, C.C., Lu, X., Chance, K., Thorpe, A.K., Duren, R.M., Miller, C.E., Frankenberg, C., Randles, C.A. 2019. Potential of next-generation imaging spectrometers to detect and quantify methane point sources from space. Atmospheric Measurement Techniques, DOI: https://doi.org/10.5194/amt-12-5655-2019.

12. Hsu Y.-K., VanCuren, T., Park, S., Jakober, C., Herner, J., FitzGibbon, M., Blake, D. R., Parrish, D. D. 2010. Methane emissions inventory verification in southern California. Atmospheric Environment., 44, 1 – 7, DOI: https://doi.org/10.1016/j.atmosenv.2009.10.002.

13. Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L. 2012. On the Sources of Methane to the Los Angeles Atmosphere. Environmental Science & Technology., 46 (17), 9282-9289, DOI: https://doi.org/10.1021/es301138y.

14. Peischl, J., Ryerson, T. B., Brioude, J., Aikin, K.C., Andrews, A. E., Atlas, E., Blake, D., Daube, B. C., de Gouw, J. A., Dlugokencky, E., Frost, G. J., Gentner, D. R., Gilman, J. B., Goldstein, A. H., Harley, R. A., Holloway, J. S., Kofler, J., Kuster, W. C., Lang, P. M., Novelli, P. C., Santoni, G. W., Trainer, M., Wofsy, S. C., Parrish, D. D. 2013. Quantifying sources of methane using light alkanes in the Los Angeles basin, California, J. Geophys. Res. Atmos., 118, 4974 – 4990, DOI: https://doi.org/10.1002/jgrd.50413.

15. Jeong, S., Hsu, Y.-K., Andrews, A. E., Bianco, L., Vaca, P., Wilczak, J. M., Fischer, M. L. 2013. A multitower measurement network estimate of California's methane emissions, J. Geophys. Res. Atmos., 118, 11,339–11,351, DOI: http://doi.org/10.1002/jgrd.50854.

16. Rafiq, T., Hopkins, F. M. Assessment of AVIRIS-NG Anthropogenic Fugitive Methane Emissions using the Vista-CA Methane Infrastructure Dataset in the Bay Area, Kern County, and the South Coast Air Basin. *In prep.*

17. Lavoie, T. N., Shepson, P. B., Gore, C. A., Stirm, B. H., Kaeser, R., Wulle, B., Lyon, D., and Rudek, J. 2017. Assessing the Methane Emissions from Natural Gas-Fired Power Plants and Oil Refineries. Environ. Sci. Technol., 51, 3373 – 3381, DOI: https://doi.org/10.1021/acs.est.6b 05531.

18. Hajny, K. D., Salmon, O. E., Rudek, J., Lyon, D. R., Stuff, A. A., Stirm, B. H., Kaeser, R., Floerchinger, C. R, Conley, S., Smith, M. L., Shepson, P. B. 2019. Observations of Methane Emissions from Natural Gas-Fired Power Plants. Environ. Sci. Technol. 53, 8976 – 8984, DOI: https://doi.org/10.1021/acs.est.9b01875.

19. Environmental Protection Agency. Facility Level Information on GreenHouse Gases Tool. 2017. https:// ghgdata.epa.gov/ghgp/main.do. (accessed 13th, October 2021).

20. California Air Resources Board. 2020. Pollution Mapping Tool Sources in Your Community. https://ww3.arb.ca.gov/ei/tools/pollution_map/pollution_map.htm. (accessed 13th, October 2021).

21. Carranza, V., Rafiq, T., Frausto-Vicencio, I., Hopkins, F. M., Verhulst, K. R., Rao, P., Duren, R. M., and Miller, C. E. 2018. Vista-LA: Mapping methane-emitting infrastructure in the Los Angeles megacity, Earth Syst. Sci. Data, 10, 653-676, DOI: https://doi.org/10.5194/essd-10-653-2018.

22. Hopkins, F.M., Rafiq, T., and Duren, R.M. 2019. Sources of Methane Emissions (Vista-CA), State of California, USA. ORNL DAAC, Oak Ridge, Tennessee, USA, DOI: https://doi.org/10.3334/ORNLDAAC/1726.

23. Rafiq, T. Duren, R. M., Thorpe, A. K., Foster, K., Patarsuk, R., Miller, C. E., Hopkins, F. M. 2020. Attribution of methane point source emissions using airborne imaging spectroscopy and the Vista-California methane infrastructure dataset. Environ. Res., Lett. 15, 1-12, DOI: https://doi.org/10.1088/1748-9326/ab9af8.

24. Thorpe, A. K., Roberts, D. A., Bradley, E. S., Funk, C. C., Dennison, P. E., Leifer, I. 2013. High resolution mapping of methane emissions from marine and terrestrial sources using a Cluster-Tuned Matched Filter technique and imaging spectrometry. Remote Sens. of Env., 134, 305 – 318, DOI: http://dx.doi.org/10.1016/j.rse.2013.03.018.

25. Thorpe, A. K., Frankenberg, C., Roberts, D. A. 2014. Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: application to AVIRIS. Atmos. Meas. Tech., 7, 491–506, DOI: https://doi.org/10.5194/amt-7-491-2014.

26. Thompson, D. R., Leifer, I., Bovensmann, H., Eastwood, M., Fladeland, M., Frankenberg, C., Gerilowski, K., Green, R. O., Kratwurst, S., Krings, T., Luna, B., and Thorpe, A. K. 2015. Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane, Atmos. Meas. Tech., 8, 4383–4397, DOI: https://doi.org/10.5194/amt-8-4383-2015.

27. Duren, R. M., Thorpe, A. K., Foster, K.T., Rafiq, T., Hopkins, F. M. Yadav, V., Bue, B. D., Thompson, D. R., Conley, S., Colombi, N.K., Frankenberg, C., McCubbin, I. B., Eastwood, M. L., Falk, M., Herner, J. D., Croes, B. E., Green, R. O., Miller, C. E. 2019. California's Methane 1 Super-Emitters. Nature 575, 180–184, DOI: https://doi.org/10.1038/s41586-019-1720-3.

28. Hook, S., Johnson, W., and Abrams, M. 2013. NASA's Hyperspectral Thermal Emission Spectrometer (HyTES). In: Thermal Infrared Remote Sensing - Sensors, Methods, Applications, Kuenzer, C. and Dech, S. (Eds.), Springer.

29. Hook, S.J., G. C. Hulley, W.R. Johnson, B. Eng, J. Mihaly, S. Chazanoff, N. Vance, Z. Staniszewski, G. Rivera, K.T. Holmes and P. Guillevic. 2016. The Hyperspectral Thermal Emission Spectrometer (HyTES) - A New Hyperspectral Thermal Infrared Airborne Imager for Earth Science, Rem. Sens. Environ., in press.

30. Hulley, G. C., Duren, R. M., Hopkins, F. M., Hook, S. J., Vance, N., Guillevic, P., Johnson, W. R., Eng, B. T., Mihaly, J. M., Jovanovic, V. M., Chazanoff, S. L., Staniszewski, Z. K., Kuai, L., Worden, J., Frankenberg, C., Rivera, G., Aubrey, A. D., Miller, C. E., Malakar, N. K., Sánchez Tomás, J. M., and Holmes, K. T. 2016. High spatial resolution imaging of methane and other trace gases with the airborne Hyperspectral Thermal Emission Spectrometer (HyTES), Atmos. Meas. Tech., 9, 2393–2408, DOI: https://doi.org/10.5194/amt-9-2393-2016.

31. Air Markets Program Data. United States Environmental Protection Agency. https://ampd.epa.gov/ampd/. (accessed 13th, October 2021).

32. Eggleston, H. S., L. Buendia, K. Miwa, T. Ngara, and K. Tanabe. 2006. IPCC Guidelines for National Greenhouse Gas Inventories. Prepared by the National Greenhouse Gas Inventories Programme, vol. 4. IGES, Japan.

33. Environmental Protection Agency. Using GHG Inventory and GHGRP Data. 2019. https://cfpub.epa.gov/ghgdata/inventoryexplorer/data_explorer_flight.html (accessed 13th, October 2021).

34. Energy Information Administration Carbon Dioxide Emissions Coefficients. 2016. https://www.eia.gov/environment/emissions/co2_vol_mass.php. (accessed 13th, October 2021).

35. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 – 2014. 2016. https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2014. (accessed 13th, October 2021).

36. Fei L., Duncan, B. N., Krotkov, N. A., Lamsal, L. N., Beirle, S., Griffin, D., McLinden, C. A., Goldberg, D. L., and Lu, Z. 2020. A methodology to constrain carbon dioxide emissions from coal-fired power plants using satellite observations of co-emitted nitrogen dioxide. Atmos. Chem. Phys., 20, 99–116, 2020, DOI: https://doi.org/10.5194/acp-20-99-2020.

37. Qu, Z., Jacob, D. J., Shen, L., Lu, X., Zhang, Y., Scarpelli, T. R., Nesser, H., Sulprizio, M. P., Maasakkers, J. D., Bloom, A. A., Worden, J. R., Parker, R. J., and Delgado, A. L. 2021. Global distribution of methane emissions: a comparative inverse analysis of observations from the TROPOMI and GOSAT satellite instruments. Atmos. Chem. Phys., 21, 14159–14175, DOI: https://doi.org/10.5194/acp-21-14159-2021.

38. Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., Blake, D. R. 2014. Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data. J. Geophys. Res.Atmos., 119, 7741–7756, DOI: https://doi.org/10.1002/2014 JD021551.

39. Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D. 2015. Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. Atmos. Chem. Phys., 15, 7049–7069, DOI: https://doi.org/10.5194/acp-15-7049-2015.

40. Varon, D. J., McKeever, J., Jervis, D., Maasakkers, J. D., Pandey, S., Houweling, S., et al. 2019. Satellite discovery of anomalously large methane point sources from oil/gas production. Geophysical Research Letters, 46, 13,507–13,516, DOI: https://doi.org/10.1029/2019GL083798.

41. Cusworth, D.H., Duren, R.M., Thorpe, Dennison, P.E., Frankenberg, C., Heckler, J., Asner, G.P., Eastwood, M., Green, R.O., Miller, C.E. 2021. Carbon dioxide emissions from power plants can be quantified globally using remote sensing, AGU Advances, DOI: https://doi.org/10.1029/2020AV000350.

42. Nassar R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., Crisp, D. 2017. Quantifying CO2 Emissions from Individual Power Plants From Space Geophysical Research Letters, DOI: https://doi.org/10.1002/2017GL074702.

43. Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., Kuester, T., Hollstein, A., Rossner, G., Chlebek, C., Straif, C., Fischer, S., Schrader, S., Storch, T., Heiden, U., Mueller, A., Bachmann, M., Mühle, H., Müller, R., Habermeyer, M., Ohndorf, A., Hill, J., Buddenbaum, H., Hostert, P., Van der Linden, S., Leitão, P., Rabe, A., Doerffer, R., Krasemann, H., Xi, H., Mauser, W., Hank, T., Locherer, M., Rast, M., Staenz, K., Sang, B. 2015. The EnMAP spaceborne imaging spectroscopy mission for earth observation. Remote Sens. 7, 8830–8857.

44. Cusworth, D. H., Jacob, D. J., Sheng, J.-X., Benmergui, J., Turner, A. J., Brandman, J., White, L., and Randles, C. A. 2018. Detecting high-emitting methane sources in oil/gas fields using satellite observations, Atmos. Chem. Phys., 18, 16885–16896, DOI: https://doi.org/10.5194/acp-18-16885-2018.

45. MethaneSAT. 2021. https://www.methanesat.org/ (accessed 13th, October 2021).

46. Article 2. Mandatory Greenhouse Gas Emissions Reporting. 2021. https://govt.westlaw.com/calregs/Browse/Home/California/CaliforniaCodeofRegulations?guid=IF29D06908B1711DF8121F57FB716B6E8&originationContext=documenttoc&transitionType=Default&contextData=(sc.Default) (accessed 18th, January 2021).

Chapter 5

Conclusion

The research presented here provides a foundation for which geospatial emissions development and analysis can be built upon for further studies of CH_4 emissions in California and beyond. This research enables critical application of geospatial datasets leveraging bottom-up and top-down estimation methods and utilizing high resolution spatial data to constrain emissions at both the facility and sub-facility scale. The three chapters outlined here improve methodologies for developing high-resolution, spatially resolved datasets for major CH_4 contributors, provide novel geospatial techniques and complimentary algorithms for source attribution, offer robust geospatial techniques for top-down source and facility emission estimation, and offer industry and sector specific emissions insight for optimizing potential mitigation strategies and policies.

In Chapter 2, this work provided a novel method for automated geospatial identification and attribution of CH_4 source emitters at close to 99% accuracy across the entire state of California using a facility and infrastructure database of close to 1 million individual features in Vista-CA. This approach systematically accounts for all anthropogenic CH_4 infrastructure within California and attempts to represent all potential sources of CH_4 emissions (except rice cultivation or other anthropogenic CH_4 emissions from soils) regardless of the expected size of emissions. Combining this detailed dataset with new, high resolution observational data of CH_4 emissions from airborne remote sensing enables a more thorough accounting of CH_4 based on actual observations that is potentially more robust than activity/emission factor methods that do not capture fugitive or anomalously large sources that are common for CH_4. Using this approach, over 500 point-source emitters collected via AVIRIS-NG were identified and attributed to infrastructure and facilities throughout California. These constituted around 40% of the total statewide CH_4 inventory and this was accomplished with low latency and high confidence outputs which ultimately has the potential to be extended to other areas as well. The advantage of this technique allows for a rapid facility-level emissions assessment that can be scalable for any given area, while the disadvantages lie in the temporal quality of data collected which is not satisfactory due to the economic and hardware constraints of performing airborne surveys. Moreover, it could potentially be difficult to replicate the Vista-CA CH_4 database in international regions as the underlying data used to engineer it might not exist in other countries.

In Chapter 3, this research developed, analyzed, and scaled both top-down and bottom-up emission estimates for Kern County and investigated their differences among publicly available bottom-up inventories. This section also confirmed the hypothesis of a larger impact on CH_4 emissions in Kern County emanating from the oil and gas sector than has been previously accounted for in publicly available CH_4 emissions inventories. In fact, the AVIRIS-NG Source Data top-down estimates from fossil sources calculated in this chapter were between 2-4 times greater than both Vista-CA bottom-up and CALGEM modeled inventories. AVIRIS-NG top-down observations also revealed potential overestimation in biogenic emissions in Kern County by a factor of 7 when compared to bottom-up inventories. Furthermore, comparing the AVIRIS-NG Source Data and Vista-CA Bottom-Up datasets to state and federal data clearly shows the stark contrast in magnitudes with the both the state and federal repositories accounting for less than 5% of the total CH_4 budget of Kern County. These clear gaps in accounting in these inventories along with differences in the modeled data compared to the observed data lends credence to the need for geospatial derived facility-scale data for optimum

In Chapter 4, this work explored process-based and fugitive CH_4 emissions from the power plant sector of California using both top-down and bottom-up measurements and characterized their super-emitter activity. With the rise in both demand for electricity and expanse of natural gas usage as an energy resource, both AVIRIS-NG top-down data and Vista-CA Bottom-Up data calculated CH_4 emissions for over 250 power plants across California and exemplified super-emitter behavior likely due to fugitive activities. Top-down estimates revealed that fugitive CH_4 emissions constituted 90% of total observed power plant emissions while modeled emissions were 30-50 times smaller than top-down observations. Furthermore, geospatial and aerial imagery analysis visibly demonstrated fugitive and process-based CH_4 emissions from individual components and infrastructure located within these power plants. State and federal repositories significantly underestimated power plant fugitive emissions when compared to the top-down observations. This chapters serves to show the evident discrepancy in public CH_4 emissions accounting in a specific sector while also necessitating further utilization in onsite emissions tracking technologies such as continuous emissions monitoring devices (CEMs) with a combination of persistent facility-level airborne emission surveys for effective and actionable mitigation of both process-based and fugitive emissions.

The associated deliverables from this research are as follows: Vista-CA CH_4 Sectoral Dataset (L1C) (shapefile, kmz),[1, 2] Vista GSAAM algorithm (python code/model builder),[1] AVIRIS-NG GSAAM attributed dataset (shapefile, kmz), AVIRIS-NG source data emission estimation methodology, AVIRIS-NG attributed facility emissions dataset (shapefile, kmz), Kern County Vista-CA facility emissions dataset (L2) (shapefile, kmz), IPCC Tier 2 Vista-CA power plant emissions dataset (L2) (shapefile, kmz), HyTES emission isolation algorithm (python code/model builder), HyTES GSAAM attributed facility dataset (shapefile, kmz).

Overall, this work provides the first step towards large-scale geospatial CH_4 emissions accounting resolved at a facility-level and can lead to enhancements in potential real time facility emissions monitoring down the line. Idealistically, a real-time or near-real time geospatial facility-scale map product could provide significant information and will undoubtedly require further development in automation and machine learning technologies.[3, 4, 5, 6] While this research offers some improvement which under a global context might not hold as much significance, the potential implications this research has is greater within a local context. For example, having quantified CH_4 emissions at sub-meter to facility-level scales enables policy makers and enforcement agencies to conduct more efficient evaluations and audits of potential violators that greatly affect local residing populations through co-emitted pollutants that may be emitted alongside CH_4. This is especially the case as CH_4 emissions from nearby facilities have significant potential health impacts as seen with the Aliso Canyon CH_4 leak in late 2015.[7] Having state-level data that can scale down to the facility level supports increased local awareness of GHG emissions and enables easier investigation. This research has demonstrated that various sectors of California's CH_4 budget have been under accounted which potentially signifies that there are facilities that are producing more than what publicly available data is stating. This could negatively impact California's evaluation metrics and further delay necessary climate change mitigation milestones and implementations. Depending on the area, sector, and facility, this work has provided estimates that are 20-60% greater than current estimates in modeled inventories especially as they relate to more intensive CH_4 producing facilities such as power plants and oil and gas infrastructure. The techniques presented here could ideally be implemented by a third-party or independent environmental auditors at scale meaning the "self-reporting" abilities of facilities that produce CH_4 and other GHG's would be removed in order to curb the skirting and inaccurate reporting of emissions. This is important as these datasets are communicated, consumed, and published by state and federal modeled inventories to be used in critical policy decision making to curb CH_4 emissions and combat climate change.

I. References

1. Rafiq, T. Duren, R. M., Thorpe, A. K., Foster, K., Patarsuk, R., Miller, C. E., Hopkins, F. M. 2020. Attribution of methane point source emissions using airborne imaging spectroscopy and the Vista-California methane infrastructure dataset. Environ. Res., Lett. 15, 1-12, DOI: https://doi.org/10.1088/1748-9326/ab9af8.

2. Hopkins, F.M., Rafiq, T., and Duren, R.M. 2019. Sources of Methane Emissions (Vista-CA), State of California, USA. ORNL DAAC, Oak Ridge, Tennessee, USA, DOI: https://doi.org/10.3334/ORNLDAAC/1726.

3. Duren, R. M., Thorpe, A. K., Foster, K.T., Rafiq, T., Hopkins, F. M. Yadav, V., Bue, B. D., Thompson, D. R., Conley, S., Colombi, N.K., Frankenberg, C., McCubbin, I. B., Eastwood, M. L., Falk, M., Herner, J. D., Croes, B. E., Green, R. O., Miller, C. E. 2019. California's Methane 1 Super-Emitters. Nature 575, 180–184. https://doi.org/10.1038/s41586-019-1720-3.

4. Carbon Mapper. 2022. https://carbonmapper.org/data/ (accessed 6th, September 2022).

5. Cusworth, D. H., Jacob, D. J., Sheng, J.-X., Benmergui, J., Turner, A. J., Brandman, J., White, L., and Randles, C. A. 2018. Detecting high-emitting methane sources in oil/gas fields using satellite observations, Atmos. Chem. Phys., 18, 16885–16896, DOI: https://doi.org/10.5194/acp-18-16885-2018.

6. MethaneSAT. 2022. https://www.methanesat.org/ (accessed 6th, September 2022).

7. California Air Resources Board. Aliso Canyon Natural Gas Leak. 2022. https://ww2.arb.ca.gov/our-work/programs/aliso-canyon-natural-gas-leak (accessed 6th September 2022).