nature

https://doi.org/10.1038/s41586-019-1720-3

Supplementary information

California's methane super-emitters

In the format provided by the authors and unedited

Riley M. Duren, Andrew K. Thorpe, Kelsey T. Foster, Talha Rafiq, Francesca M. Hopkins, Vineet Yadav, Brian D. Bue, David R. Thompson, Stephen Conley, Nadia K. Colombi, Christian Frankenberg, Ian B. McCubbin, Michael L. Eastwood, Matthias Falk, Jorn D. Herner, Bart E. Croes, Robert O. Green & Charles E. Miller

Supplementary Information for

California's Methane Super-Emitters

Riley M. Duren[1,2], Andrew K. Thorpe[1], Kelsey T. Foster[1], Talha Rafiq[3], Francesca M. Hopkins[3], Vineet Yadav[1], Brian D. Bue[1], David R. Thompson[1], Stephen Conley[4], Nadia K. Colombi[5], Christian Frankenberg[6,1], Ian B. McCubbin[1], Michael L. Eastwood[1], Matthias Falk[7], Jorn D. Herner[7], Bart E. Croes[7], Robert O. Green[1], Charles E. Miller[1]

[1,*]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.[2]University of Arizona, Tucson, AZ, USA. [3]University of California Riverside, Riverside, CA, USA. [4]Scientific Aviation, Boulder, CA, USA. [5]University of California Los Angeles, Los Angeles, CA, USA, [6]California Institute of Technology, Pasadena, CA, USA, [7]California Air Resources Board, Sacramento, CA, USA

*Corresponding author: Riley.M.Duren@jpl.nasa.gov

Material and Methods

S1.1 Survey design

The next generation Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-NG) measures ground-reflected solar radiation from the visible to infrared spectral regions (380 to 2,510 nm) with 5 nm sampling. This push broom instrument has a 34° field of view and operates on high performance aircraft, allowing for efficient mapping of large regions (5). Increasing flight altitude affects the ground resolution, i.e., the size of each image pixel increases while the image swath increases. For most of this survey, AVIRIS-NG flew at 3 km above ground level, resulting in 3 m image pixels. The airborne survey program with AVIRIS-NG (Fig. 1) was conducted with intensive field campaigns in 2016, 2017 and 2018 – with the longest single campaign lasting over 2 months. Each flight day was typically 4-5 hours in duration including cruise flight to the mapping area. Mapping was typically conducted between the hours of 1000 and 1500 local time for peak illumination but some flights occurred as early as 0900 and as late as 1700.

Figure S.1 illustrates an example of AVIRIS-NG flight lines in the Southern San Joaquin Valley including the diversity of emission sectors and their spatial distributions. The flight planning was governed by two primary objectives: 1) spatial coverage sufficient to map the infrastructure in the State most likely responsible for >60% of methane point source emissions (with >80% coverage for key sectors) and 2) sufficient number of revisits to have a reasonable probability of detecting intermittent emission sources (e.g., for a source that is active 25% of the time, 6 visits should provide a detection probability of 0.82). In addition to the broader goal to map and revisit large areas we also conducted several intensive studies focused on gaining insight into key emission processes. One intensive focused on an area near Visalia that was mapped repeatedly over a 5 hour period to investigate the temporal variability of manure emissions from over 100 dairies with 60 minute revisit intervals. Others focused on natural gas infrastructure across southern California, gas-fired power plants during heat wave conditions and refineries in the LA basin and San Francisco Bay Area. In several cases coordinated, contemporaneous measurements were conducted with mobile onroad laboratories, fixed surface

observations and other airborne systems to help validate source locations and emission estimates (Section S3).

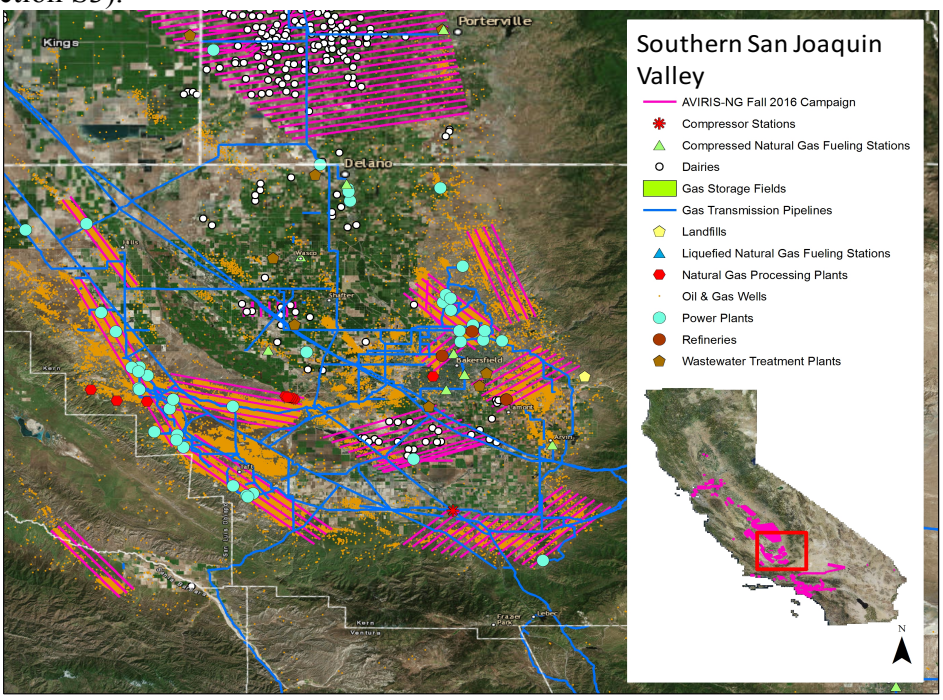

Figure S.1. Example of flight planning for the southern San Joaquin Valley in Fall 2016 using the Vista-CA GIS data system and the resulting AVIRIS-NG flight lines. Surface map data: Google Earth.

S1.2 Infrastructure Data

A Geographic Information System (GIS) data set known as Vista-CA that maps potential methane emitting infrastructure across the State of California was developed to assist with flight planning and for source attribution following detection of methane plumes. The Vista-CA data set applies similar methods previously used to develop a Vista-LA methane GIS data set for the greater Los Angeles area (32). Vista-CA mapped the locations of infrastructure associated with three primary sectors (energy, agriculture, and waste) following the frameworks used by the State of California's Greenhouse Gas Inventory and the IPCC Guidelines for GHG Reporting. Vista-CA contains 450,572 distinct pieces of potential methane emitting infrastructure and was used to guide selection of flight boxes (Table S.1).

Many of the Vista-CA elements were readily derived from public data records but others were more challenging and required some new development. For example, the natural gas pipeline numbers in Vista-CA include transmission, distribution, gathering and "other" categories (Table S.2). The 4,599 km of gas transmission lines in California was derived from a combination of NMPS, CEC and EIA data but there is no publicly available map of distribution lines in urban areas. To the distribution line issue we constructed a residential distribution line mask using parts of the California road network overlaid on raster cells classified as being 20-100% impervious in urban areas from the National Land Cover Dataset (NLCD; ref 33) and connected it to the existing NG Pipeline infrastructure using a 10km distance tolerance. Survey coverage was computed by using the AVIRIS-NG 2016-2017 flight path (1800m width)

rectangular polygons as clip features to pull out overlapping pipelines and recalculate segment length within each AVIRIS survey polygon.

Table S.1. Summary of Vista-CA infrastructure elements, data sources and % surveyed by this study for key equipment/facility types and IPCC emission sectors.

IPCC Source Category	Vista-CA Infrastructure Elements		Vista-CA Infrastructure Elements Data Sources	Number of Features Surveyed by AVIRIS (2016-2017)	Total Number of Vista-CA Infrastructure Elements	Percentage of Vista-CA Infrastructure Elements Surveyed	Percentage of IPCC Emission Sector Surveyed
1A1 Energy Sectors	Gas-fired Power Plants		CARB Inventory (2014) EIA (2016) EPA FLIGHT (2016)	238	435	54.7%	57.3%
	Refineries		CARB Inventory (2014) EIA (2016) EPA FLIGHT (2016)	26	26	100.0%	
	sub-totals			*264*	*461*		
1B2 Oil and Natural Gas	CNG Fueling Stations		AFDC (2017)	107	162	66.0%	60.4%
	LNG Fueling Stations		AFDC (2017)	25	46	54.3%	
	Natural Gas Stations (non-storage compressor, dehydration, metering, odor, etc)		CEC (2017) EPA FLIGHT (2016)	538	1,131	47.6%	
	Natural Gas Pipelines (length in km)		CEC (2012) EIA (2017) NLCD (2011) NPMS (2013) U.S. Census Bureau (2017)	68,548	216,774	31.6%	
	Natural Gas Processing Plants		EIA (2014)	23	26	88.5%	
	Natural Gas Storage Fields		DOGGR (2016) EIA (2016)	12	12	100.0%	
	Oil and Gas: Other production equipment		DOGGR (2018)	2,872	3,356	85.6%	
	Oil and Gas: Wells		DOGGR (2018)	198,231	225,766	87.8%	
	sub-totals			*270,356*	*447,273*		
3A2 Manure Management	Dairies	All dairies	CIWQS (2018) CARB (2015) RWSCB - Region 5 (2017) SJVAPCD (2017)	890	1,544	57.6%	64.5%
		CAFOs with >1000 head		443	620	71.5%	
4A1 Solid Waste Disposal Sites	Landfills	Composting Sites	CalRecycle (2015) CARB (2015) EPA FLIGHT (2016)	166	430	38.6%	38.2%
		Solid Waste Disposal Sites (landfills)	CalRecycle (2015) CARB (2015) EPA FLIGHT (2016)	270	716	37.7%	
4D1 & 4D2 Wastewater Treatment	Wastewater Treatment Plants	Domestic & Industrial Wastewater treatment	CARB (2016) EPA FLIGHT (2016)	57	148	38.5%	38.5%
		Industrial Wastewater treatment: beef processing	other (satellite imagery)	1	n/a	n/a	
		TOTALS		***272,447***	***451,192***	***60.4%***	

"Natural Gas Stations" in Table S.1 include 158 non-storage compressor stations as well as dehydration stations, metering stations, odor stations, pressure limiting stations, regulation stations, storage stations, taps, and valves. The "Oil and Gas: Wells" category includes active well heads, pumpjacks, and other equipment immediately associated with extraction and also inactive wells. "Oil and Gas: Other Production Equipment" is derived the "California Statewide Oil and Gas Production or Injection Facility Boundary" data set from DOGGR https://maps.conservation.ca.gov/doggr/metadata/FacilityBoundaries.html. To our knowledge that is the best publicly available database on locations of oil and gas production infrastructure in California that may emit methane including permanent tanks, flowlines, headers, gathering lines, wellheads, heater treaters, pumps, valves, compressors, injection equipment, production safety systems, separators, manifolds, and pipelines. However, that database does not likely cover all such equipment statewide, and also excludes production equipment known to emit methane, such as separators, water tanks, acid gas removal units, and dehydrators. We have very limited information about the spatial distribution of some components such as gathering lines. For these reasons (and the fact that we surveyed over 80% of production fields in the State) we do not attempt to upscale our emissions results from Other Production Equipment (Section S2.11).

Table S.2. Details on key elements within the Vista-CA natural gas pipeline layers.

Data Source/Designation	Network Type	Total length (km)	Surveyed by AVIRIS (km)
CA Roads/NLCD Impervious Surface generated Network	Distribution	196,670	62,184
NPMS/CEC/EIA Network	Distribution	1,311	565
	Transmission	4,599	1,842
	Gathering	0	0
	Other	14,193	3,958
TOTALS		216,774	68,548

Dairies are a special case given the number and magnitude of methane sources and complexity in identifying which facilities are more likely to be point source emitters. We determined the locations of dairy farms in California primarily from satellite imagery viewed in Google Earth. Dairy farms have a distinctive appearance in aerial imagery, typically identifiable by the presence of metal-topped shelters, manure lagoons, and corral areas located together in close proximity. This method identified 1,709 individual dairy facilities. 855 of these farms were overflown during AVIRIS-NG flight campaigns, yielding 58% coverage of dairies statewide.

California has experienced a contraction in the number of dairies over time, namely a reduction by 30% from 2002 to 2012 (34). Hence it is possible that some of the infrastructure from now-closed facilities may still be standing and would be included in the aerial imagery identification. To validate the number of dairies that are still in operation, we used a list of permitted facilities from the California Integrated Water Quality System (CIQWS) Project (e.g., National Pollutant Discharge Elimination System surface water discharge permits required under the Clean Water Act). CIQWS provides the facility name and location for all active, permitted dairies in the state (35). Out of the 1,709 dairies identified from aerial imagery, 1,172 were linked to CIQWS permitted facilities. 296 additional facilities were identified by other permit

data (36,37), for a total of 1,468 permitted dairies, which are assumed to be currently operational. Among the active facilities, 855 were overflown by AVIRIS-NG (58% coverage).

Not all dairies are created equal with respect to methane emissions; there is large variability in the size and manure management practices across California (37). To isolate the dairies most likely to emit measurable amounts of methane from manure management, we chose to focus on dairies with >1,000 milk cows, which are expected to emit ~26 kg CH_4 per hour, roughly 3 times the detection limit for AVIRIS-NG. This criteria is similar to the Environmental Protection Agency (EPA)'s definition of a large concentrated animal feeding operation (CAFO), which is ≥1,000 cattle. The 2012 USDA Census reported 642 dairy CAFOs (facilities with >1,000 milk cows) operating in California (38).

To determine which dairies were CAFOs, we used records obtained from the California Regional Water Quality Control Board (RWQCB) Central Valley (Region 5) and Santa Ana (Region 8) offices for 2015 (36). These annual reports contain the name, address, and herd size for all dairy farms located within these two regions. These regions account for more than 95% of milk production in the state (39), 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 (37). The counties encompass 846 of the 855 active dairies flown by AVIRIS-NG, leaving 9 dairy overflights from the counties of Imperial (8 facilities) and Monterey (1 facility), where herd size reporting is not currently required.

In total, we had 825 records for dairies in the Central Valley, and 78 for dairies in the Santa Ana region reporting more than 1 milk cow for 2015. Of these 903 dairy reports, 383 were large CAFOs, with > 1,000 milk cows (355 in the Central Valley, and 28 in the Santa Ana Region). Among these large CAFOs, 282 were overflown by AVIRIS-NG (270 in the Central Valley, and 12 in the Santa Ana Region), yielding 74% coverage.

There is uncertainty associated with the estimate of CAFOs, specifically from (a) dairies in counties that are not required to report herd numbers to the RWQCBs (9 flown dairies), (b) dairies from the original dataset that did not have RWQCB reports despite being in the reporting regions (179 flown dairies), and (c) possible mis-reporting of herd sizes in RWQCB reported data.

For (a), we attempted to quantify the number of CAFOs for counties not included in RWQCB reporting using a combination of USDA data and CIQWS number of farms at the count level. According to the USDA 2015 survey of agriculture counts of dairy cows, the counties that had >1,000 cows were Del Norte, Humboldt, Imperial, Marin, Mendocino, Monterey, San Diego, San Luis Obispo, Santa Barbara, and Sonoma (38). We estimated the maximum number of CAFOs in each county by using the minimum of either (1) the 2015 USDA county cow count divided by 1,000 (Del Norte: 2, Marin: 3, San Luis Obispo: 0, San Mateo: 0); (2) the 2012 USDA census of agriculture number of cows on dairies with >500 head divided by the number of farms in the CIQWS dataset (Humboldt: 4, Sonoma: 16, San Diego: 2); or (3) when data was withheld from USDA surveys to avoid disclosing data for individual operations, we assumed that the maximum number of CAFOs was equal to the number of CIQWS farms (Mendocino: 2, Monterey: 1, Santa Barbara: 1) or was equal to 3, following the threshold rule (Imperial: 3; ref 51). That yielded a maximum of 34 CAFOs outside the RWQCB regions described above. This increases the total number of CAFOs in the state by 34, and the total number of CAFOs flown by 4, as only dairies in Imperial and Monterey counties are included.

For (b) we estimated the number of CAFOs missing from the RWQCB dataset by assuming the proportion of CAFOs among the missing farms was the same as that in the dataset by county (e.g., in Madera county, CAFOs were 50% of the dataset, so for the 18 farms were excluded from the dataset, 9 were considered to be CAFOs, and of the 2 flown dairies missing herd data, 1 was considered to be a CAFO). This increased the total number of CAFOs by 194 farms, and the number flown by 153 farms.

Finally, for (c) there is likely error associated with reported herd numbers in the RWQCB permit data. According to RWQCB there may be 15% more cows than permitted on farm at any given time[1]. To account for this, we estimated how much our CAFO estimate would change given milk cow populations at each farm that were 15% greater (or less) than reported. With 15% greater (lesser) herds, the number of CAFOs would increase by 65 (decrease by 63), and the number of flown CAFOs would increase by 48 (decrease by 46).

Adding in the uncertainty from (a) and (b) as corrections to our dataset, together with uncertainty from (c), we estimate a total of 620 CAFOs in the state (with a possible range of 557-685), and 443 were flown (397-491). This yields 71% coverage, or a scalar of 1.40 (1.13-1.73) that is used for the sectoral upscaling described in section S2.11.

S2.1 Analysis workflow

The analysis for this study (Fig S.2) consists of a) standard processing including calibration and orthorectification of the AVIRIS-NG image cube data, b) retrieval of methane column mixing ratio-lengths and generation of methane plume maps, c) quality control and filtering of plumes, d) geolocation and attribution of methane plumes to Vista-CA spatial layers, e) calculation of integrated methane enhancement (IME) and length for each plume, f) acquisition and processing of High Resolution Rapid Refresh (HRRR) reanalysis wind fields, g) emission flux estimation and uncertainty quantification for individual methane plumes, h) filtering and removal of plumes that exhibit sub-optimal shapes, are redundant/overlapping with others plumes or have excessive errors in IME and/or wind speed estimates, i) validating emission estimates with independent methods, j) averaging and scaling plume emission estimates with observed persistence to derive an annual net emission for each source, k) applying Vista-CA spatial layers to calculate net emission estimates for facilities and key sectors statewide, l) apply bootstrap analysis to determine confidence intervals for each sector and total population. Each of these steps is described below.

[1]https://www.waterboards.ca.gov/centralvalley/board_decisions/adopted_orders/general_orders/r5-2007-0035.pdf

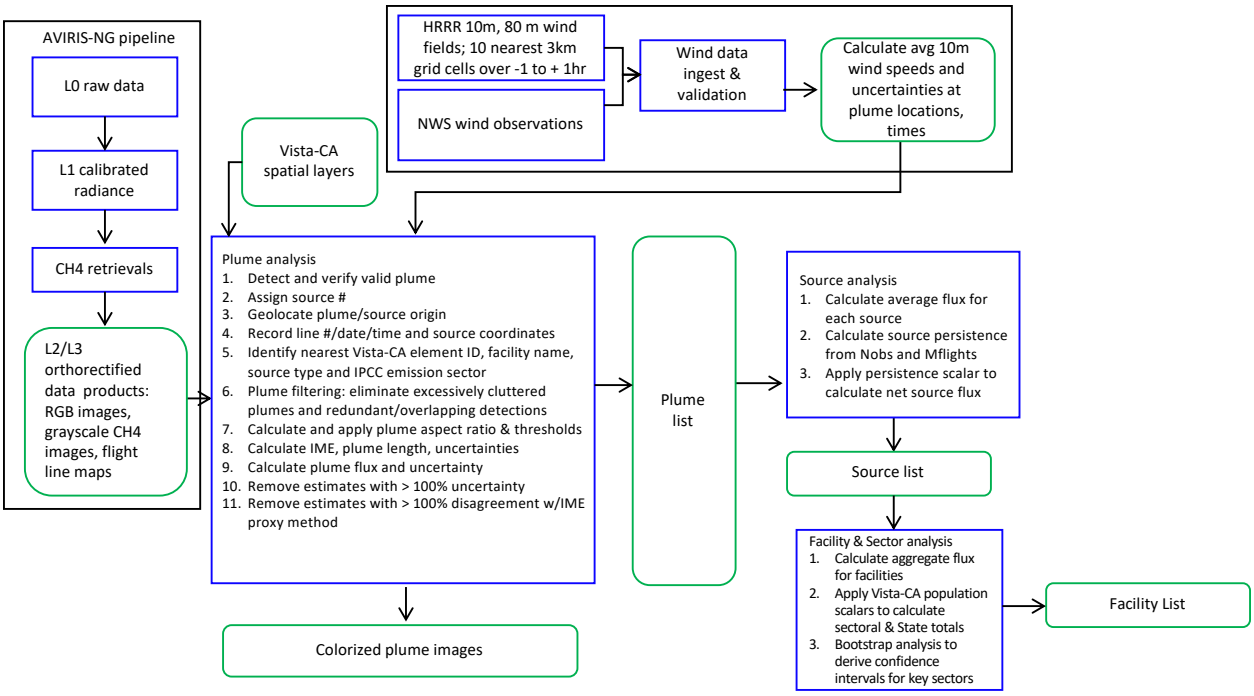

Fig S.2 Analysis workflow for this study.

S2.2 Standard image processing and methane retrievals

The AVIRIS-NG standard data pipeline acquires "level 0" (L0) raw image data cubes from the instrument and delivers calibrated and orthorectified radiances to be analyzed for atmospheric methane. The methane retrieval is based on absorption spectroscopy between 2,100 and 2,500 nm and uses a linearized matched filter to estimate α, a mixing ratio length that represents a methane enhancement in parts per million meter (ppm-m). The retrieval algorithms have been previously described (7,26) and demonstrated in multiple previous field campaigns including oil fields in California (7), a campaign to the Four Corners region in Colorado and New Mexico (9), the Aliso Canyon gas leak (26), and a study of California landfills (24). Controlled release experiments with natural gas demonstrated reliable detection of methane plumes for emission rates as low as 2-10 kgCH_4/hr depending on AVIRIS-NG flight altitudes, surface brightness and wind speeds (41). The retrieved quantity is equivalent to an excess methane concentration (or mixing ratio) in ppm if the layer is one meter thick (i.e. directly equivalent to ppb km). At a scale height of about 8 km, the total column averaged excess mixing ratio would be about 0.000125 times the excess in ppm-m. For example, 1000 ppm-m is equivalent to a column enhancement of 125 ppb.

S2.3 Source detection, location and attribution

The AVIRIS-NG "level 3" (L3) data products include the as-flown nadir flight-line tracks as well as the native Red, Green, Blue (RGB) images of surface features and grayscale images of retrieved methane along each flight line. Human analysts analyze the grayscale methane images to detect and geolocate plumes. An experimental machine learning system – a convolutional neural network- was also trained on a subset of plumes from this and other field studies and then used to assess potential false positives and false negatives in the manual plume list. The origin

of each plume – i.e., the emission source - is estimated based on plume shape and enhancement values. Quality control analysis is performed on every observed plume, including visual inspection to evaluate plume shape, presence of noisy retrieval results, instrument artifacts, correlation of plume direction with surface wind direction (for single and multiple detections), and plume proximity to known surface infrastructure (attribution). For the latter step, we apply the instrument RGB imagery and higher resolution satellite imagery as well as our Vista-CA GIS system. Confidence in the source location and attribution to infrastructure increases in cases where plumes are repeatedly detected over multiple flights.

S2.4 Plume Filtering

At this stage all plumes are counted as valid sources in terms of threshold detection however noisier plumes are excluded from the emission estimation steps. We start by establishing a mixing ratio threshold of 1,000 ppm-m. We also define a plume aspect ratio metric A as the measured plume length in number of pixels divided by the plume area in number of pixels. For plumes with A in the closed interval [0.02,1.0] we use the threshold of 1,000 ppm-m and for all others we apply a 1,500 ppm-m threshold. Based on empirical evaluation this scheme rejects plumes that are difficult to interpret – particularly those with high methane background levels or excessively non-gaussian in shape - without eliminating smaller plumes.

S2.5 IME and Plume length calculation and uncertainties

We isolate each observed plume by applying the mixing ratio threshold and two parameters: maximum fetch (radius in meters from the plume origin) and merge distance which allows for definition of contiguous plumes in the presence of gaps (e.g., pixels with low methane values). For this analysis maximum fetch was set to 150 meters and merge distance was set to 20 meters based on iterative assessment of optimal plume size and shape for emission estimation (see Figure S.3 A).

Centered on the plume origin an initial circle of radius (r_1) is defined based on the pixel size (Figure S.3 B). For the portion of the plume covered by this circle, the total excess mass of methane in the plume that we refer to as Integrated Methane Enhancement (*IME*) is calculated as the summation of the methane mixing ratio length α for all pixels above the detection threshold multiplied by the pixel area S and converted to CH_4 mass units with the constant k (7,26). As shown in Figure S.3 B, the radius is sequentially expanded (r_c, where c denotes the total number of circles) and new IME for the given radius (IME_{r_c}) is calculated as follows:

$$IME_{r_c} = k \sum_{i=0}^{n} \alpha(i) S(i)$$

Next, the IME for the given radius is divided by the radius of the circle (r_c; i.e., the plume length) which results in a ratio of IME_{r_c}/r_c. To minimize variability in these calculations associated with plume shape this ratio is calculated for the entire range of radii (r_c) and an average and standard deviation is calculated, representing the uncertainty in the *IME/r* estimates, $\sigma_{IME/r}$. The mean fractional error $\frac{\sigma_{IME/r}}{\overline{IME/r}}$ was 29% for the population of plumes used in the analysis.

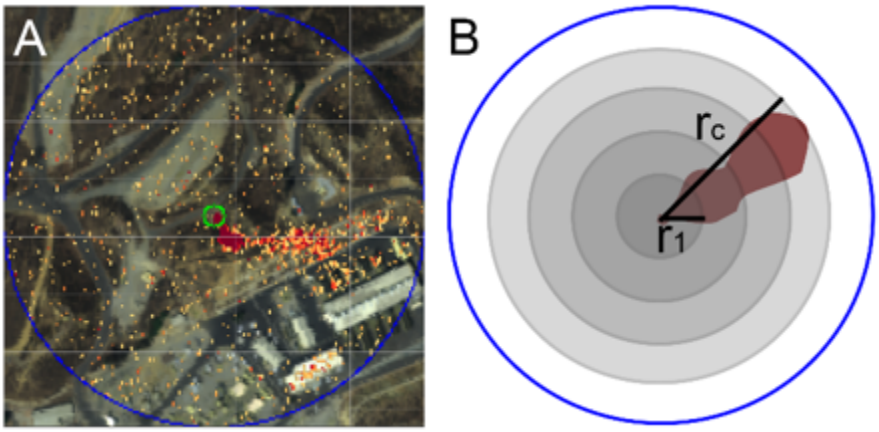

Figure S.3 A: Example methane plume with maximum fetch radius (blue circle) with respect to the plume origin (green circle). B: Key parameters for calculating IME_{r_c}/r_c for the entire range of radii (r_c). Surface map data: AVIRIS-NG.

S2.6 Wind speed, turbulence and uncertainties

Estimating emission rates requires information about the near-surface wind speed. Surface weather observations are sparse over the large area covered by this study so our analysis uses spatially continuous wind fields from NOAA's High-Resolution Rapid Refresh (HRRR) product for the coterminous United States (42). We use HRRRv3 10m wind fields in forecast mode, available as an hourly, 3km grid product from the HRRR archives at U. Utah (43) and NOAA (44).

To account for spatial and temporal variability in HRRR wind fields for each of the methane plumes used in this analysis we compute an average wind speed at 10 meters above ground level (U_{10}) from a total of 27 HRRR grid cells: the 9 nearest to the plume origin (e.g., a 3x3 box centered on the source) each repeated for 3 time-steps (plume detection time $\pm$ 1 hour). For the entire population of methane plumes detected in this study, U_{10} ranges from 0.4 to 9.2 m s^{-1} (mean 3.0m s^{-1}). For wind speed uncertainty we compute σ_u as the standard deviation of the 27 HRRR grid cells for each plume detection, ranging from 0.04 to 1.92 m s^{-1} (mean 0.44 m s^{-1}). In terms of relative wind speed error $\frac{\sigma_U}{U_{10}}$ these translate to 1% to 69% (mean 22%).

To account for the possibility of wind shear resulting in significantly lower wind speeds at the surface we used the following standard logarithmic relationship to derive atmospheric stability (45):

$$U(z) = U_{ref}\left(\frac{Z}{Z_{ref}}\right)^{\varepsilon}$$

where $U(z)$ is the wind speed at the target height z above ground level, U_{ref} is the wind speed at the reference height Z_{ref} above ground level and ε is the shear (atmospheric stability) exponent. HRRR provides wind speeds at 10m and 80m so we solved for ε at every detected plume location and time. For our entire population of methane plumes the atmospheric stability ε

ranges from to 0.45 (strongly stable) to -0.18 (strongly unstable) with a mean 0.04 (moderately unstable). Nearly 90% of the methane plumes in this study occur in unstable conditions which translates to minimal wind shear. This is expected given that AVIRIS-NG flights are typically conducted between the hours of 1000 and 1500 during which insolation and turbulence are at a peak. This – and the fact that many methane plumes originate from elevated infrastructure and that large eddy simulations indicate significant lofting except under high wind conditions - provides confidence that additional vertical correction is not required.

To evaluate our wind speed data, we compared hourly averaged National Weather Service (NWS) surface wind observations (46) at 130 airports distributed across California with overlapping HRRR 10 m wind fields for 70 mid-day periods distributed across Fall 2016 and Fall 2017. Comparison showed no evidence of overall bias; the mean wind speed of NWS observations was 0.1 m s^{-1} higher than HRRR, lower than the mean HRRR uncertainty of 0.44 m s^{-1} and consisted with a similar analysis by Varon et al 2018 (47). The correlation between HRRR and surface observations varies from an R^2 of 0.20 to 0.93 (mean 0.70) for each 4 hour mid-day window during the 70 day study. The RMS error of the NWS surface observations vs the HRRR 10 m wind speeds for the entire data set ranges from 0.79 to 3.42 m s^{-1} (mean 1.9 m s^{-1}), consistent with a 1.6 m s^{-1} RMS error reported by Varon et al comparing GEOS-FP reanalysis (10m) wind fields and Mesonet weather stations across the US (47).

Most of the methane plumes detected in this survey have ventilation times < 500 seconds and can thus be influenced by variability in wind speeds that occur on time-scale shorter than our hourly averaged HRRR-derived estimates. To evaluate this possibility we also acquired 5-minute observations from the same NWS stations for an entire year (October 2017 – September 2018). We found that our HRRR wind estimates report wind speeds that are 50% lower on average than the NWS 5-minute observations. This indicates that some of our methane emission estimates in this study are potentially conservative under-estimates. Additionally, we acquired surface wind observations from a meteorological tower at 2m and 10m above ground level with 5 minute averaging and NCEP Real Time Mesoscale Analysis (RTMA) wind fields at 10m, 15 minute averaging over a 5 hour period (48). For the same location and times we also derived a 3 hour, nearest 10 grid cell average from the latter to simulate our standard method that uses HRRR hourly average wind speeds. The 2m and 10m wind speeds derived from the observations and RTMA generally agreed to within 10-20% on 5 to 15 minute time-scales, within the mean uncertainty bounds for the HRRR based wind speeds for our population of methane plumes.

Turbulence and large eddies are not significant sources of uncertainty in our study because a) we're using the Integrated Methane Enhancement (IME) method to estimate emissions which is intrinsically less sensitive to turbulent diffusion at the finest scales and b) our high spatial resolution (3 meters for this study) resolves the near-field plume response including turbulence and eddies within the 150 meter maximum fetch. Plume variability due to turbulence and large eddies at scales larger than 150m does not affect our emission estimates. Varon *et al.* (47) refer to the inadequacy of gaussian plume modeling for addressing spatially small methane plumes because they depart from gaussian behavior due to turbulence. Varon *et al.* also highlights the strong variability due to turbulent diffusion when using the source pixel estimation method. It is for these reasons that we use the Integrated Methane Enhancement (IME) estimation method. Varon *et al.* explicitly highlight the advantages of the IME method: "We show that standard methods to infer source rates by Gaussian plume inversion or source pixel mass balance are prone to large errors because the turbulence cannot be properly parameterized

on the small scale of instantaneous methane plumes. The integrated mass enhancement (IME) method, which relates total plume mass to source rate, and the cross- sectional flux method, which infers source rate from fluxes across plume transects, are better adapted to the problem". The IME approach has been previously demonstrated to be robust with independent in-situ measurements and the cross-sectional flux method (9).

S2.7 Source emission estimation and uncertainties

We calculate the emissions for each plume (Q, in $kgCH_4\ hr^{-1}$) using the excess mass of methane in the plume (the average $\overline{IME}$ in kg), the average plume length r in m and near surface wind speed (U_{10}):

$$Q = \left(\overline{IME}/_{r}\right) U_{10}$$

The IME method is inherently less sensitive to errors in wind speed than Gaussian plume inversion or single pixel mass balance and cross-sectional estimation methods given the extended nature of the plumes and the information content from hundreds to thousands of pixels (47). We confirmed that the uncertainties associated with $\overline{IME}/_{r}$ and U_{10} are uncorrelated for the set of ~1000 methane plumes used in the analysis and hence the relative errors σ_{IME} and σ_U combine in quadrature to provide the total uncertainty σ_Q for each emission estimate:

$$\sigma_Q = Q\left(\left(\frac{\sigma_{\overline{IME}/_{r}}}{\overline{IME}/_{r}}\right)^2 + \left(\frac{\sigma_U}{U_{10}}\right)^2\right)^{\frac{1}{2}}$$

We obtain individual source uncertainties ranging from ±4 to ±95% (mean 30%); see Fig S.5 for distribution of source emission estimates and uncertainties. We eliminate from calculation of source emissions the small fraction of plume estimates that exhibit >100% uncertainty (approximately 20 plumes). Performance with our framework is consistent with the theoretical best-case performance of 15-50% uncertainty for an equivalent precision instrument and ideal plumes predicted by large eddy simulations (47). We consider these levels of uncertainty to be acceptable for this analysis given the very sparse data regarding point sources in California many of which have never been identified, much less precisely geolocated or quantified with uncertainties. For example for the Aliso Canyon gas blowout study, in situ methane sampling with aircraft using Gauss's theorem resulted in 1σ uncertainties ranging from ±9 to ±22% (23). With appropriate meteorological conditions, uncertainties below ±10% are possible with that method down to detection threshold of about 5 kg hr^{-1} (49,54). That technique is limited to estimating the net emissions from a facility (not spatially explicit) and typically requires prior knowledge of source location. Another studying applying a mass balance approach with SWIR retrievals of column averaged methane from the Methane Airborne Mapper (MAMAP) instrument for California landfills resulted in uncertainties in the ±14 to 45% range (24). Limitations with that technique include the time required to perform many flight-line transects over a source and the lack of imagery to help pinpoint source locations.

S2.8 Source filtering and accounting

Due to the close proximity of some sources, multiple plumes can fall within the same search radius used to determine plume length (see Figure S.4). If this occurs, multiple emissions estimates are generated for the same or similar spatial extents leading to plume overlap and an overestimate of emissions. To prevent this, plumes with high amounts of overlap were removed from the emissions analysis. The potential area within which a plume can be detected is represented as a circle with a 150m radius from the source origin which is the same as the maximum fetch value used in IME calculations. For each flight line, the percent overlap between potential plume area circles is calculated. If there is $\geq 30\%$ overlap then the plume with the highest percent overlap is removed and percent overlap is recalculated. This process is repeated until the maximum percent overlap is below 30%. See figure S.4 for a visual of this process.

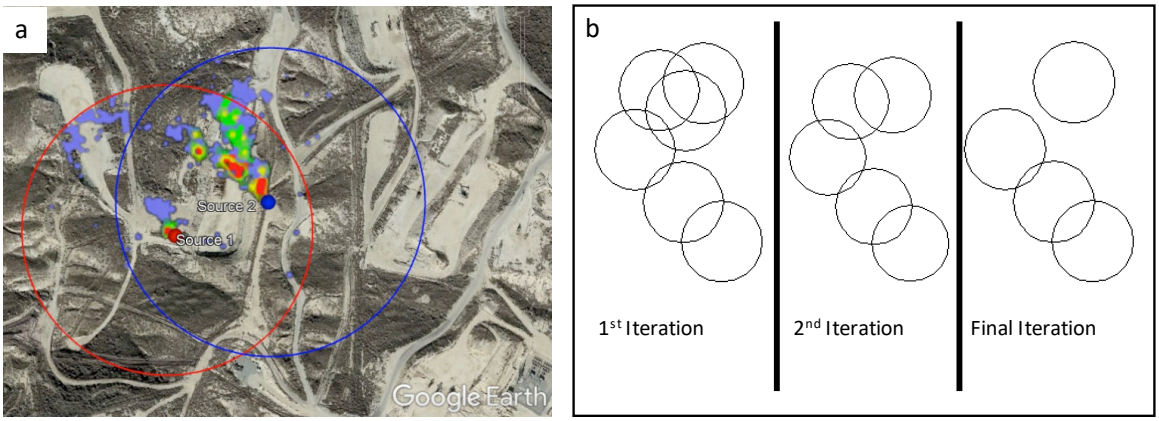

Figure S.4. (a) Example image showing two overlapping plume detection areas. The red and blue points indicate the location of plumes and the red and blue circles indicate the extent in which a plume can be detected. The plume search radius used for Source 1 and Source 2 overlap resulting in overlapping plumes. (b) Potential area a plume can be detected in represented as circles. Circles are iteratively removed until percent overlap between circles is less than 30%. Surface map data: Google Earth.

For most sectors, the extent of the observed methane plume was small compared to the full spatial extent of the associated facility and generally appeared in a repeatable fashion from the source to which it was attributed. Hence for most sectors we report emissions for individual sources, with larger facilities often including multiple sources. However, we apply a different accounting scheme for landfills given the complexity of emission processes. For landfills were plumes were detected we observed large plumes that spanned the spatial extent of the facility. Additionally, in most cases the location of each landfill plume evolved significantly over time in response to daily changes in waste deposition and surface cover. We therefore defined each landfill with observed methane plumes as a composite source. All plume observations at a given landfill, within a single flight line, were summed to get the total facility emissions per flight line for that sample interval. In cases where multiple flight lines were required to cover the whole facility, these flight lines were combined into a composite flight line. Flight lines were

only combined into a composite flight line if those flight lines were from the same day, were less than 16 minutes apart (typical time for a complete landfill survey), and had less than or equal to 50% overlap. Plumes identified within composite flight lines were then summed, with duplicate plumes removed by excluding the plume with the lower emissions estimate. There were some instances where an emission estimate was not generated for a plume (see previous description of filtering and quality. control). If an individual or composite flight line had more than one plume without an emission estimate or spatially covered less than 85% of a facility, then the flight line was removed from the analysis. The remaining individual and composite flight lines were averaged by facility to obtain source estimates for each landfill.

Since the focus of this study is characterizing a broad population of emitters rather than individual sources, as a further check on the robustness of plume emission estimates derived with our standard method we apply the IME proxy method described below in Section S.3 to generate a second set of emission estimates that does not rely on wind information. Plume emissions that exhibit > 100% difference between our standard estimation method and IME scaling method are flagged as outliers and removed from further analysis. This step removed an additional 170 plumes from the analysis but did not eliminate sources.

The net result of all of the above filtering steps resulted in 1050 plumes with emission estimates that were attributed to 564 sources (Fig S.5). Compared to no filtering, these steps conservatively reduce the number of sources with emission estimates by 20% and the net emissions by 7%.

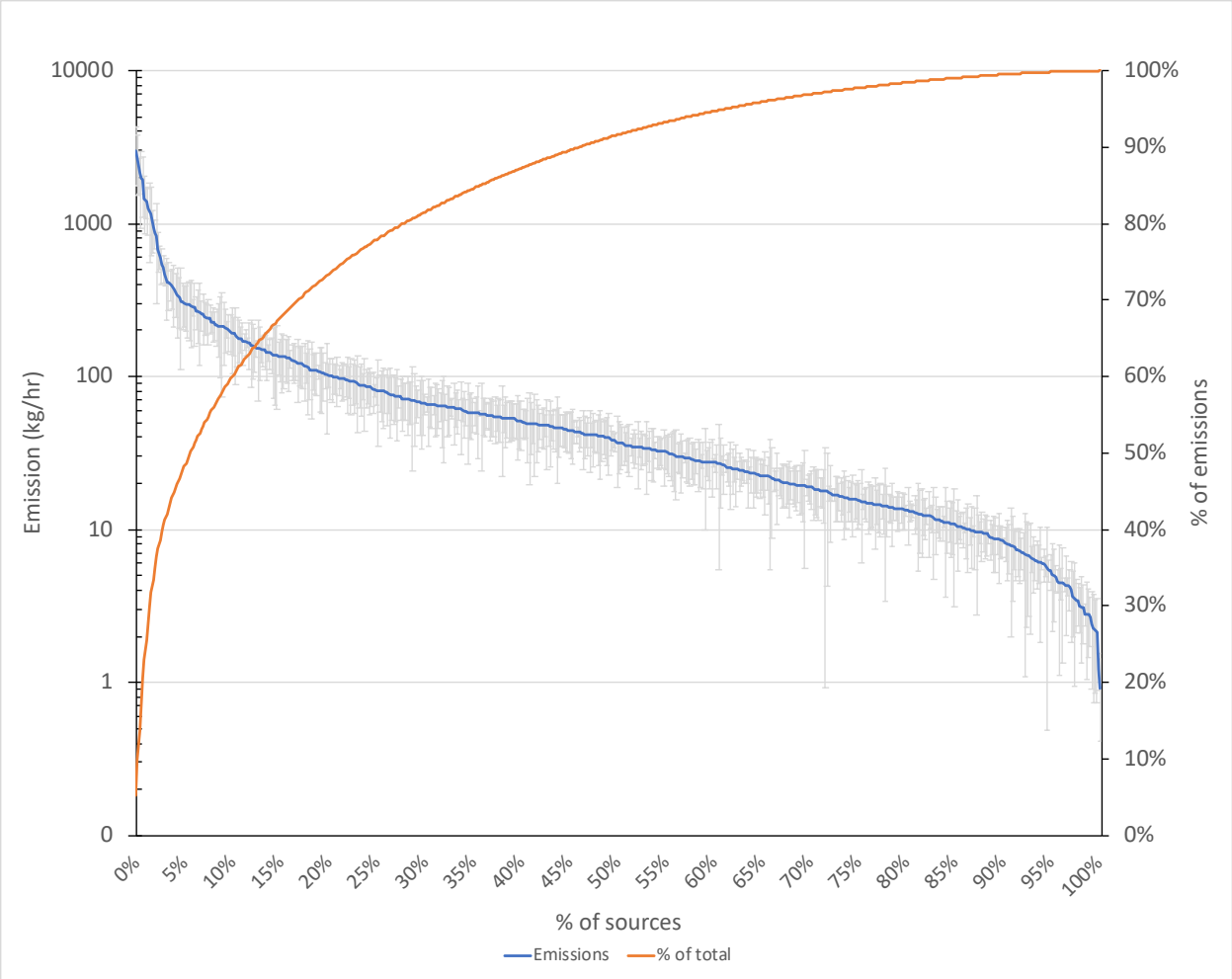

Fig. S.5: Distribution of persistence-adjusted average methane emissions and 1σ uncertainties (1 s.d.) for 564 point sources. The heavy-tail indicates that 10% of detected point sources contribute 60% of the total population emissions.

S2.9 Persistence calculation and adjusted emissions

Nearly every methane source in this study was overflown by AVIRIS-NG multiple times. For each source we calculate an average emission rate $\overline{Q}$ from all plume emission estimates that pass our filtering criteria. We also calculate the persistence or frequency of emissions f of a given source in order to account for intermittency (where a source may not emit above detection threshold during other overflights). f is the ratio of the number of observed plumes M to the total number of overflights N for a given source.

$$f = \frac{M}{N}$$

We define N as the number of flight lines where a source intersects the area covered by an overflight. AVIRIS-NG flight outline files from https://avirisng.jpl.nasa.gov/alt_locator/ are used to represent the area covered by flight lines. Due to orthocorrection errors some flight

outlines were not generated. For those flight lines where a flight outline is not generated a theoretical flight outline calculated as the product of the theoretical swath width and the flight line length. The flight line length is the distance between the start and stop coordinates of a flight line and theoretical swath width was calculated as:

$$2\mathrm{H}\tan\left(\frac{FOV}{2}\right)$$
$$where\ H = altitude\ and\ FOV = field\ of\ view$$

We then calculate the total emission rate Q_t for each source using the average emission $\overline{Q}$ (from good plume emissions estimates) scaled by the source persistence *f:*

$$Q_t = f\overline{Q}$$

The persistence metric provides important information about source activity independent of quantitative emissions. Determining the persistence for a source population also provides an opportunity to assess the detection confidence for a given observing system. The probability *P* of detecting a source with persistence ≥ *f* after *N* overflights (for emissions ≥ the detection threshold) is

$$P = 1 - (1 - f)^N$$

For example, for a source with a 33% persistence (the mean for our source population) there is a 96% probability of detection with 8 overflights. We calculate *P* for every persistence estimate in our source list (Section S4), resulting in a median of 92% (mean 82%) over the total population.

S2.10 Potential for diurnal bias

Diurnal variability of selected methane emission processes can in some cases introduce temporal sampling bias. This possibility exists for nearly every top-down (atmospheric) emission estimation method, including most aircraft measurement techniques (in-situ and remote-sensing) as well as regional scale flux inversions using measurements from satellites and regional in-situ monitoring networks. With the exception of portable flux chamber measurements (limited to a very small number of sites), most top-down methods rely on an understanding of mixing and ventilation processes that are often poorly represented in atmospheric transport models – particularly at night. Hence most top-down studies typically focus on sampling during midday when the planetary boundary layer is well mixed. For our study, sampling of solar reflectance was conducted between the hours of 0900 and 1700 local time with most samples occurring between 1000 and 1500, including weekdays and weekends.

Alvarez et al (16) also addressed this topic and pointed out that there could actually be higher emission rates for some processes (e.g., abnormal activity not detected and remedied by operators) during off-shift hours. Eliminating the possibility of diurnal sampling bias for a source population as large as this study would require an unprecedented investment in near-continuous daytime and nocturnal sampling over large areas over several months. Nocturnal remote sensing of methane emissions is possible with imaging spectrometers operating at

thermal wavelengths and/or active systems (lidar) both of which introduce additional challenges that would likely preclude the kind of spatial coverage achieved in this study.

We discuss here the potential magnitude of any such bias for key sectors. Regarding the potential for diurnal bias from mid-afternoon (1300-1600) aircraft in-situ mass-balance surveys of oil and gas production, Vaughn *et al* (50) conducted a study of the Fayetteville oil and gas basin for a two month period in Fall 2015, including two days of activity data provided by operators. They reported a ~ 37% increase in basin scale methane emissions attributed to manual liquid unloadings with peaks near the time of the afternoon overflights. In that case half of the basin scale emissions were attributed to time-varying activity with other contributors in production, gathering and transmission activity reported to be non time-varying. However, unloading operations are generally limited to unassociated gas wells which in California are concentrated primarily in the Sacramento Valley which was only responsible for 1% of our observed methane oil and gas emissions. Additionally, Alvarez et al (16) also did a probabilistic analysis that concluded that large (>10,000 $kgCH_4\ h^{-1}$) maintenance related blowdowns of high pressure gas systems is only likely to occur during 1-10% of daytime hours. Hence, episodic events such as manual unloadings and blowdowns are unlikely to contribute a significant daytime bias for sampling oil and gas emissions in California.

With regards to potential diurnal variability from dairy methane point sources, Wood *et al* (51) studied 6 manure lagoons over a 6 month period and reported median increases of 5-20% for discrete sampling between 1000-1600 local time compared to continuous sampling. Taylor *et al* (52) reported a roughly +/- 25% deviation in landfill methane emissions from the daily mean for mid-day measurements. The actual variability in methane point-sources at dairies and landfills was not determined by Wood *et al* and Taylor *et al* because those studies address the net facility emissions including point- and area-sources.

Based on the aforementioned studies of diurnal variability our assessment is that the potential impact of temporal sampling bias could be as large as +/-25% for some facilities (roughly equivalent to the +/-30% mean uncertainty we report for individual emission estimates) but is likely somewhat smaller. The true impact of diurnal variability is difficult to determine because these studies of diurnal bias themselves carry significant uncertainty in terms of their spatial completeness (a much sparse set of facility scale assessments than this study) and they do not attempt to disentangle the relative contributions of area-source and point-source mechanisms to net diurnal variability. We underscore the need for additional, continuous observational studies of methane sources globally to characterize the true diurnal variability of emissions and potential magnitude of any temporal sample bias.

S2.11 Aggregating emissions and uncertainties to facilities, key sectors and statewide

To estimate total emissions for facilities we use the Vista-CA spatial layers to determine the locations and boundaries of facilities and infrastructure elements (Section S.1). We then sum the persistence adjusted average emissions from all detected point sources that fall within those boundaries. As described in section S2.8 for landfills this aggregation to facility scale occurs at the source accounting level given the complexities with landfill emissions.

Additionally, since we surveyed between 32% and 100% of every point source emission type in California we can also upscale our measured emissions to estimate the complete distribution of all point source emissions in the State for the key sectors and infrastructure types covered by this study. To do this we derive scalars for each sector or infrastructure type from the

total number of Vista-CA elements and the number of those elements observed at least once by this study (Table 1 and Table S1). For example, Vista-CA contains 435 gas-fired power plants and we surveyed 238 of them so that sectoral scalar (435/238=1.83) is used to convert our measured emissions from that sector from 0.07 to a state point source total for that sector of 0.013 $TgCH_4\ y^{-1}$. This approach was used for all entries in Table 1 with three exceptions – all of which result in smaller scaled emissions. First, Vista-CA's representation of infrastructure elements in the oil and gas production sector is incomplete. While we include all known 225,766 oil and gas production sites (wells and collocated equipment) in California the remainder of production equipment elements in Vista-CA is not well quantified because of gaps in public data regarding other, non-extraction equipment in each facility. Vista-CA includes the outlines of 3,356 non-extraction site production equipment in California such as condensate tanks and we can calculate how many of those elements were surveyed (2,872). However we have very limited information about the number and locations of gathering lines in oil and gas fields. Deriving a scalar for the Other Production Equipment sector would require making assumptions about the prevalence of those equipment types and so we conservatively do not attempt to upscale that sector in this analysis (scalar in Table 1 is set to 1.0). We do however provide a scalar for the Production Sites (well pads, pump-jacks, and co-located equipment). Second, while we surveyed 430 out of 1,127 landfills and composting facilities statewide the methane emissions our experience is that point source only appear at a small fraction of these facilities and so applying a simple scalar could over-estimate the emissions from this sector. Instead, we used findings from a process model based analysis of California landfill emissions (53) to prioritize our survey of this sector. We prioritized 430 landfills in this study based on predictions that those facilities should be responsible for ~90% of the managed waste disposal total point source methane emissions. Hence, our scalar for this sector is only 1/0.9 = 1.11. Third, while we did not explicitly include industrial wastewater treatment and discharge in the Vista-CA data set or our flight planning we detected a single large beef processing facility with methane plumes emanating from slurry pits. After confirming the latter were not associated with dairies or other nearby infrastructure (using satellite imagery) we allocated this single facility to emission sector 4D2 "Industrial Wastewater: beef processing".

As described in Section S1 we estimate a total of 620 Confined Animal Feeding Operations (CAFOs) in the state, each with at least 1,000 head of cattle (with a possible range of 557-685). Since 443 (397-491) were surveyed in this study this yields 71% coverage, or a scalar of 1.40 (1.13-1.73). This scalar is somewhat smaller than if we had used all dairies covered by our survey. Our imaging data indicates that methane point sources associated with manure management is primarily limited to wet management (e.g., anaerobic lagoons) at CAFOs and not dry management at smaller dairies and beef cattle feedlots.

Finally, given that the observed source emissions are not normally distributed (based on the Anderson Darling test for normality) and given the significant variability between sectors we applied a non-parametric bootstrap analysis to estimate confidence intervals for the emission estimates for six cases: the 5 IPCC sectors covered by this study as well as the total source population. For each case we resampled the observed distribution of emissions with replacement and collected 10,000 bootstrap replications of the sum of emissions. The resulting envelope of emissions was used to compute 95% confidence intervals (Table S.3.). Those confidence intervals along with the central value of emission totals for each case where then multiplied by the Vista-CA scalars (Table 1) to estimate statewide total emissions and confidence intervals. We make two assumptions in this upscaling. One, for the cases of the Energy Industries and Oil

and Gas sectors we assume that the confidence interval is consistent across each sector and so apply them equally to every sub-sector when applying the Vista-CA scalars (e.g., for the Energy Industries sector we assume that the 95% confidence interval of -45/+57% of sector total emissions applies equally to emissions from gas fired power plants and refineries, respectively). The risk of this assumption is mitigated by the fact that the sub-sectors dominating each sector's emissions also represent the largest number of sources in the sector so outliers should have a small impact on the total. Two, when upscaling to address unobserved infrastructure we assume that there is no spatial bias in the Vista-CA scalars. Our assessment is this is a low risk assumption given that each of our measured sectoral populations are well distributed across the key infrastructure centers and climate zones in the state – including Northern California, the Central Valley and Southern California – and hence already accounted for by our empirical bootstrap analysis.

Table S.3 Sum of measured source emissions and 95% confidence interval (2.5 and 97.5% limits) for each sector and the total population derived by bootstrap analysis.

	Oil and Gas		**Managed Waste Disposal**		**Manure Management**		**Energy Industries**		**Waste Water Treatment**		**Total**	
Sum of emissions from sample population (kgCH4/hr)	15614		26179		13125		2546		509		58388	
	lower	upper	lower	upper	lower	upper	lower	upper	lower	upper	lower	upper
Percentile CI (kgCH/hr)	13328	18117	17973	35492	11189	15280	1395	3989	229	876	49416	68568

S2.12. Distribution of sectoral emissions and comparison with CARB inventory

We find a similar distribution of emissions in the manure management, waste water treatment and oil and gas sectors (Fig. S.6). However, for the managed waste disposal sector we only detected methane plumes at 32 out of 436 surveyed facilities. Landfill methane emissions observed in this study are qualitatively different in terms of their 100% persistence compared to the 20-35% (mean) persistence for point sources in the other sectors. The observed landfill emitters include some of the largest outliers in our overall source population and collectively are the highest emitting point source sector in California (Fig. S.6).

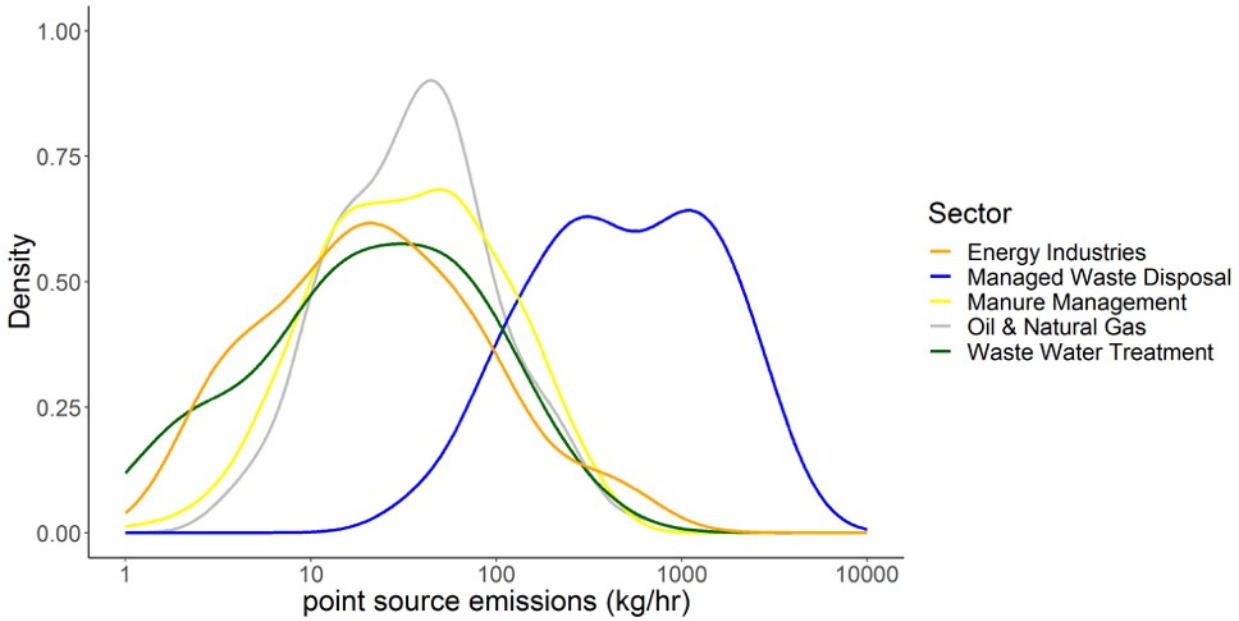

Fig. S.6 Histograms indicating the density of measured methane point source emissions (adjusted for persistence) for each of the key sectors in California ($kgCH_4\ h^{-1}$). Managed waste disposal exhibits qualitatively different behavior than the other sectors, with point sources only appearing at 32 persistent, high emitting landfills – likely constituting a distinct sub-population within that sector.

Our study provides insights into the relative contributions of methane point sources to the relevant emission sectors in CARB's methane inventory (Fig. S.7), the latter representing 67% of the State total. The point source emissions in our study are most likely limited to these sectors rather than others that are most likely dominated by area sources (e.g., rice cultivation and enteric fermentation). Given that our net emissions are equivalent to 34-46% of the State total inventory this suggests that point sources may constitute an even greater percentage of these sectors (e.g., 1/0.67 or 51-69%) or that those sector totals are significantly under-estimated in the CARB inventory (e.g., the contribution from area sources could be larger than understood). As a minimum, Figure S.7 indicates significant inventory under-estimates for the managed waste disposal and energy industries.

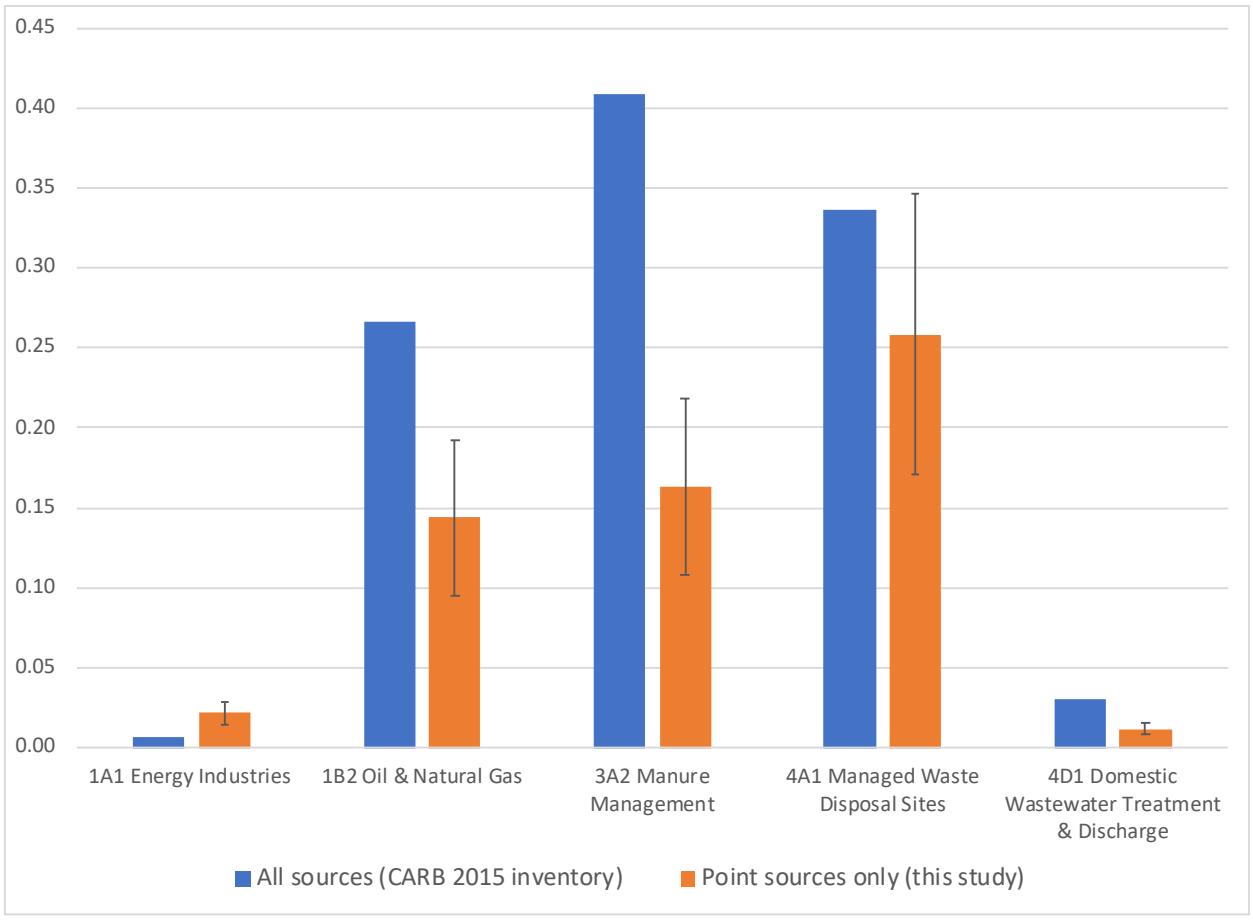

Figure S.7 Comparing statewide methane point source emission estimates from this study and the relevant sectors in the CARB inventory that are likely to include point sources (8). The whiskers indicate the 95% confidence intervals from this study.

S2.13 Emissions from the waste management sector

As discussed in the main text, methane point sources were only detected at 32 landfills (including 2 identified as composting facilities) which collectively represent a distinct sub-population from the rest of managed waste disposal sector. We attribute some these persistent outliers to anomalous behavior with landfill cover and gas capture systems based on high spatial resolution imaging of plumes at the outlier facilities. Figure S.8 illustrates the ability of this method to detect methane plumes from landfills at high resolution. The AVIRIS-NG overflight for Figure S.8 occurred following installation of new landfill gas transport features at a large landfill prior to completion of new gas capture systems. The three circles indicate the locations of those features, confirmed by the operator. We shared our data collected over this facility from 2016-2017 with the landfill operator who used it to help guide efforts to reduce emissions.

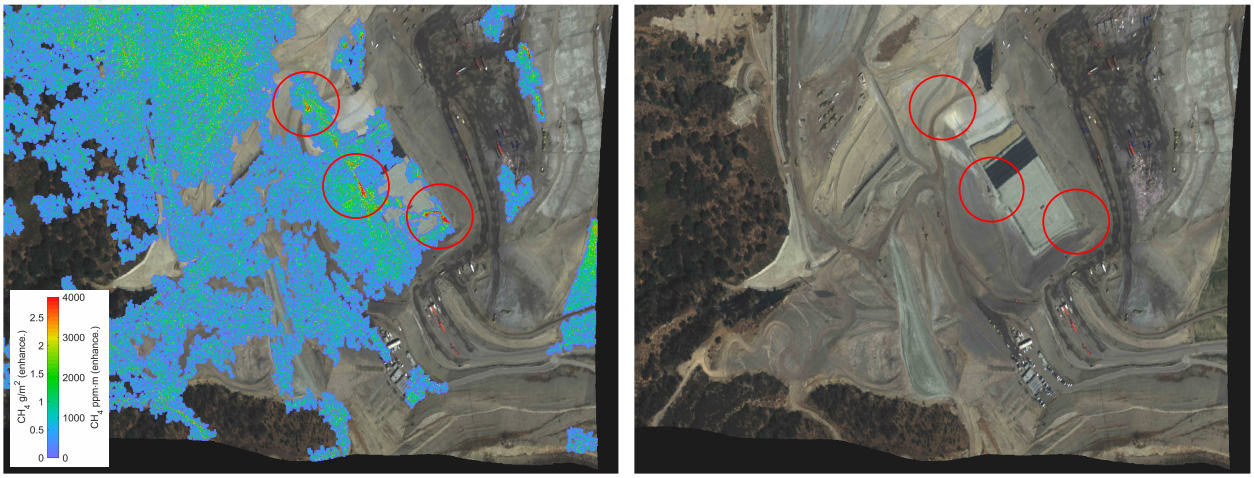

Figure S.8 Example of landfill methane emissions due to installation of structures for transporting landfill gas to the surface, ultimately for gas capture. Capture systems were not yet in place at the time of this overflight. Surface map data: AVIRIS-NG image.

S2.14 Emissions from refineries and power plants

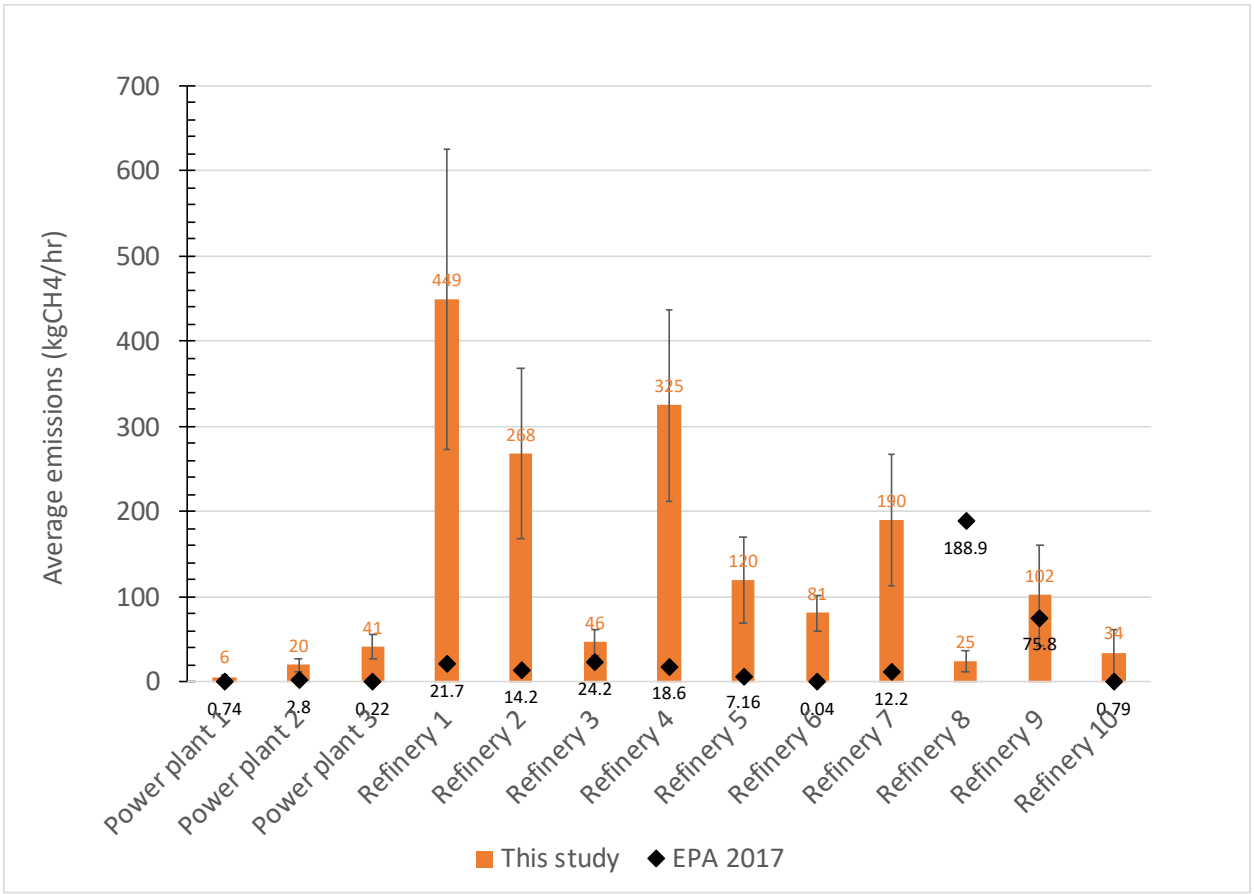

Figure S.9 Persistence-adjusted average emissions estimated by this study compared with 2017 GHGRP reported emissions (28) for refineries and power plants. With two exceptions the GHGRP emissions are significantly lower than observed by AVIRIS-NG. The errors bars indicate 1 s.d.

S3. Validating emission estimates

S3.1. Comparison with independent observations

Methane retrievals and IME estimation methods were validated in previous studies (7,26). Detection limits and emission estimation methods were previously demonstrated in field studies (7,9, 26,41) and with large eddy simulations (47). However, since our study is the first to attempt to estimate emissions for individual methane sources from a large population distributed across a wide region over multiple years we also undertook additional validation steps. We conducted simultaneous observations of several facilities in Fall 2017 as well as additional non-simultaneous follow-up flights of those and other facilities that spanned the dynamic range of emission rates with an independent airborne measurement system (Fig S.10; 31). Scientific Aviation has a well-established method of collecting in-situ measurements of methane and other trace gas mixing ratios and wind speeds aloft with a Mooney aircraft (23). The system applies Gauss's theorem to derive robust mass-balance emission estimates with a detection threshold comparable to AVIRIS-NG (~ 5kg h^{-1}) and typically source uncertainties <20% (49).

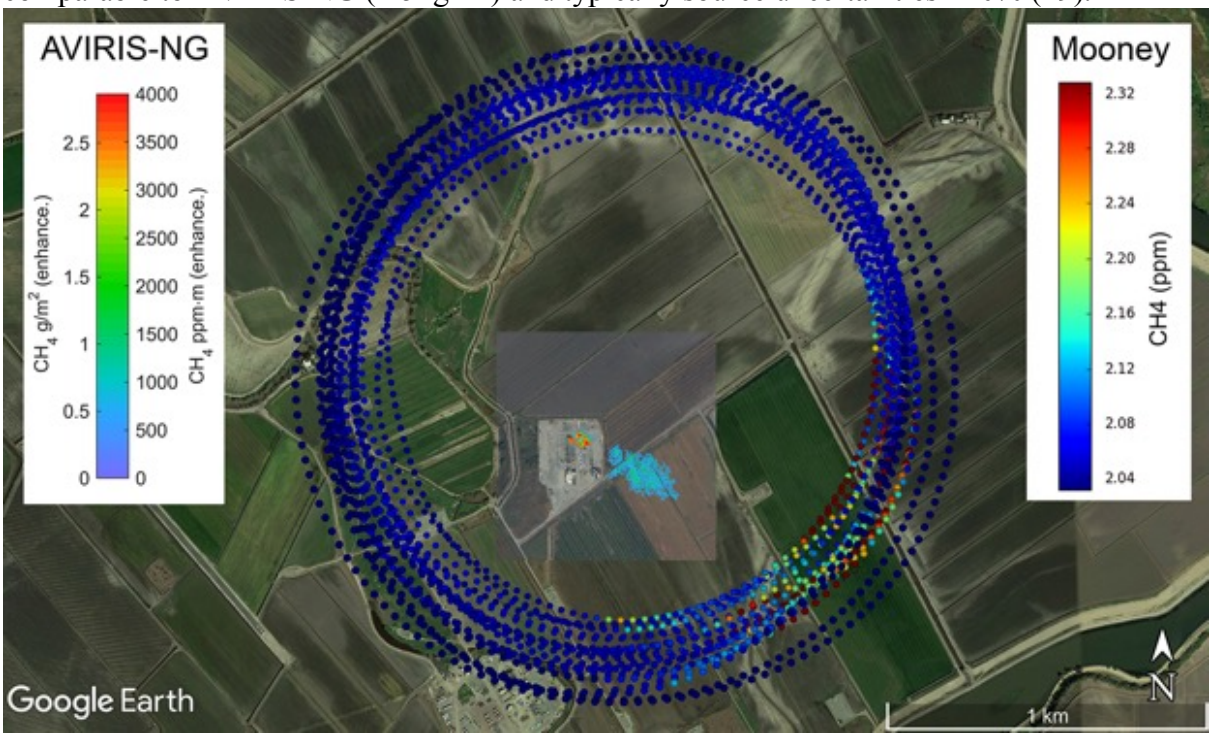

Figure S.10 Example of a coordinated survey with two aircraft for a natural gas storage facility ("gas storage facility 1" in other figures). The ~1.5 km diameter spiral pattern indicates the integrative in-situ sampling by Scientific Aviation's Mooney aircraft - typically a 30-60 minute operation for each source (methane mixing ratio measurement legend on the right). The central image of an individual gas plume from a compressor station blowdown stack is from the AVIRIS-NG imaging spectrometer. During such intensive surveys AVIRIS-NG typically conducts 3-4 over-flights at 3 km altitude, 1.8 km wide swath, 3 meter pixels with 10-20 minute revisit intervals (methane enhancement legend on the left). The AVIRIS-NG plume aligns with the higher mixing ratios in the Scientific Aviation data. Winds were out of the north west. Surface map data: Google Earth and AVIRIS-NG image.

The remote-sensing methods used in this study were previously validated against Scientific Aviation's independent in-situ mass-balance method in the Four Corners region (9). We are using the same general approach here. However, unlike the Four Corners study which focused on isolated point sources such as oil and gas production sites and coal mine vents, in this study we have observed a more diverse set of facilities including landfills and dairies. The latter are known to include both point- and area-source contributions to the total emissions for each facility (e.g., large dairies generate both point source plumes from manure management and diffuse area source emissions from enteric fermentation). This must be considered when attempting to compare results from the two methods: AVIRIS-NG can only measure the point source emissions while Scientific Aviation measures the net (combined area and point source) emissions from a given facility. The AVIRIS-NG estimate should be lower for a facility with area source emissions. We address intercomparison in two steps. First, we compare AVIRIS-NG measurements with a known emission rate from a controlled release experiment and with simultaneous Scientific Aviation observations of isolated point sources. Next, we compare AVIRIS-NG and Scientific Aviation emission estimates for 14 diverse facilities including simultaneous, single flights and averaging over multiple flights spanning several months.

Figure S.11 provides a comparison of instantaneous snapshots of sources that are as close as possible to pure point sources, albeit for emission rates below 300 $kgCH_4$/hr due to constraints on allowable venting rates and availability of reliable energy sector point sources. A controlled release test was conducted in Fall 2017 with the assistance of a natural gas operator who vented small, precisely metered quantities of gas from a pipeline in a remote location in the desert. We instrumented the site with sonic anemometers to provide high frequency near-surface wind speed and direction observations as a check on HRRR wind fields. AVIRIS-NG conducted multiple overflights of the release site at our typical survey altitude of ~ 3km while the operator changed the release rates in 3 steps from 16 to 89 $kgCH_4\ h^{-1}$(each with <5% uncertainty). Given that the controlled release test was limited to fairly small emission rates we also conducted simultaneous overflights of several energy facilities with Scientific Aviation that provided 4 isolated point sources with higher emission rates up to ~300 $kgCH_4\ h^{-1}$. In addition to confirming the degree of correlation between the known emissions and AVIRIS-NG estimates (Figure S.11a) we also used these results to derive a simple empirical scaling between the AVIRIS-NG derived IME and the known emissions rates (Figure S.11b).

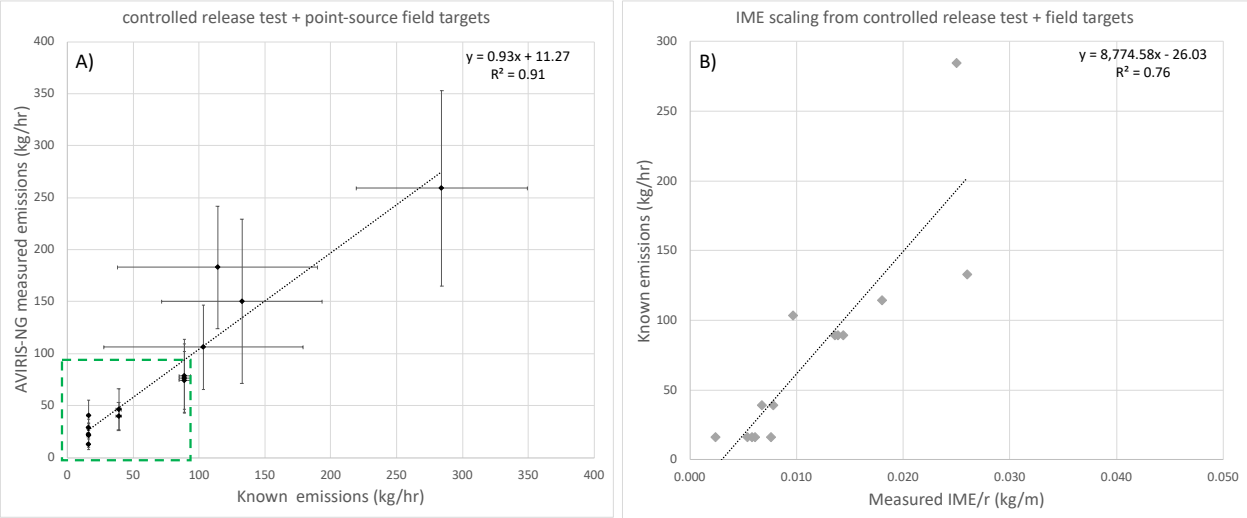

Figure S.11 A) Results of a controlled release test from a natural gas pipeline for known rates up to ~ 90 kg/hr (green box) combined with estimates for four higher rate, isolated point sources with single plumes jointly observed by AVIRIS-NG and Scientific Aviation during coordinated overflights. The controlled release test rate and Scientific Aviation estimates were collectively treated as known emissions within the indicated uncertainties. B) Corresponding correlation between observed IME/r values from AVIRIS-NG for each of the independently known emission rates. The resulting linear fit provides a method for scaling the observed IME values from other sources as a simple check on our emission estimates. The error bars represent 1 s.d.

Figure 3 and Figure S.12 provide an assessment over the broader range of emission types and time-scales. We show results from coordinated AVIRIS-NG and Scientific Aviation surveys of representative facilities spanning the range of emission rates and sectors. Figure 3a compares estimates for simultaneous, single flight observations. Figure 3b compares average emission estimates derived from multiple, non-coordinated flights of the two aircraft over several months. These estimates correlate with an R^2=0.86 for all 14 facilities and 0.99 for the 8 facilities with simultaneous surveys. The Scientific Aviation mass balance method is able to measure the net emissions from each sampled area including the combined contributions of area sources and point sources whereas the AVIRIS-NG method is only sensitive to point sources.

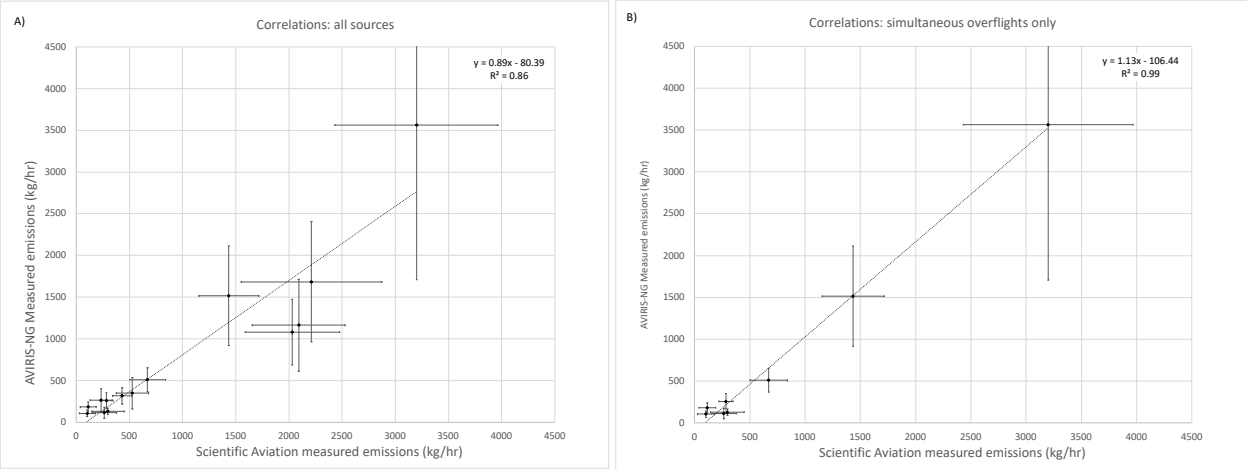

Fig. S12 Correlation plot indicating the degree of agreement between AVIRIS-NG and Scientific Aviation estimates shown in Figure 3 for A) all 14 facilities and B) for the 8 facilities with simultaneous flight lines. The error bars represent 1 s.d.

We also applied the relationship between IME and emissions from our test case derived in Fig. S11b to derive a coarse emission estimate for our entire population without using wind information. We then compared those results with our standard method that uses HRRR wind data (Figure S.13). In cases where the disagreement between our standard method and the IME proxy was > 100% we flagged the plume estimate as an outlier and it from further analysis.

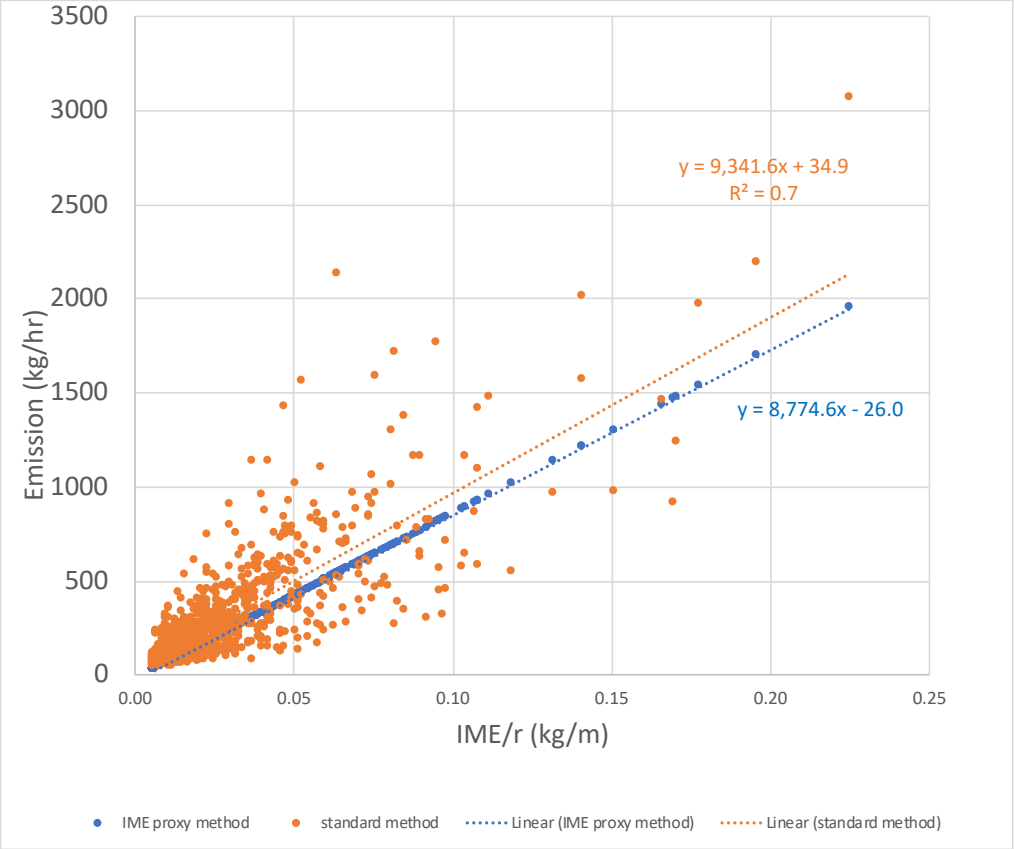

Figure S.13 Comparison of estimates for 1050 plumes using a) the standard emission analysis method used in this study (Section S2.7) and b) the IME proxy (scaling) method to derive coarse estimates. (Fig. S.11b).

S.3.2 Comparison with bottom-up data

There are currently few opportunities to compare our top-down (atmospheric measurement) estimates with bottom-up estimates of facility scale methane point source emissions because of gaps in the latter category of data sets. Greenhouse gas inventories are aggregated at coarse spatial scales and on an annual basis using standard emission factors and thus offer limited insight into the distribution of individual point source emitters. However, it is possible to do a limited comparison using self-reported emissions for selected sectors. We identified 351 individual facilities in California that participated in the EPA's GHG Reporting Program (GHGRP) in 2017 (28) - excluding a single number from SoCalGas representing fugitives across their entire service network and 36 facilities reporting zero emissions. Figure S.14 plots the distribution of those GHGRP emissions along with our emission estimates for 250 facilities. The plot indicates that 99% of the emissions from the two distributions originate from facilities each with >= 10 $kgCH_4\ h^{-1}$ for the GHGRP in California and >= 25 $kgCH_4\ h^{-1}$ for our study (the latter matches the 99% level for GHGRP emissions for the entire US). We also note that over half (0.281 $TgCH_4$) of our estimate for statewide point source emissions is allocated to dairies and oil and gas production, neither of which are included in the GHGRP. Given this and the discrepancies between reported emissions and top-down emission estimates using two independent methods (Figure 3) this further underscores the utility of remote sensing in addressing a gap in understanding facility level methane emissions.

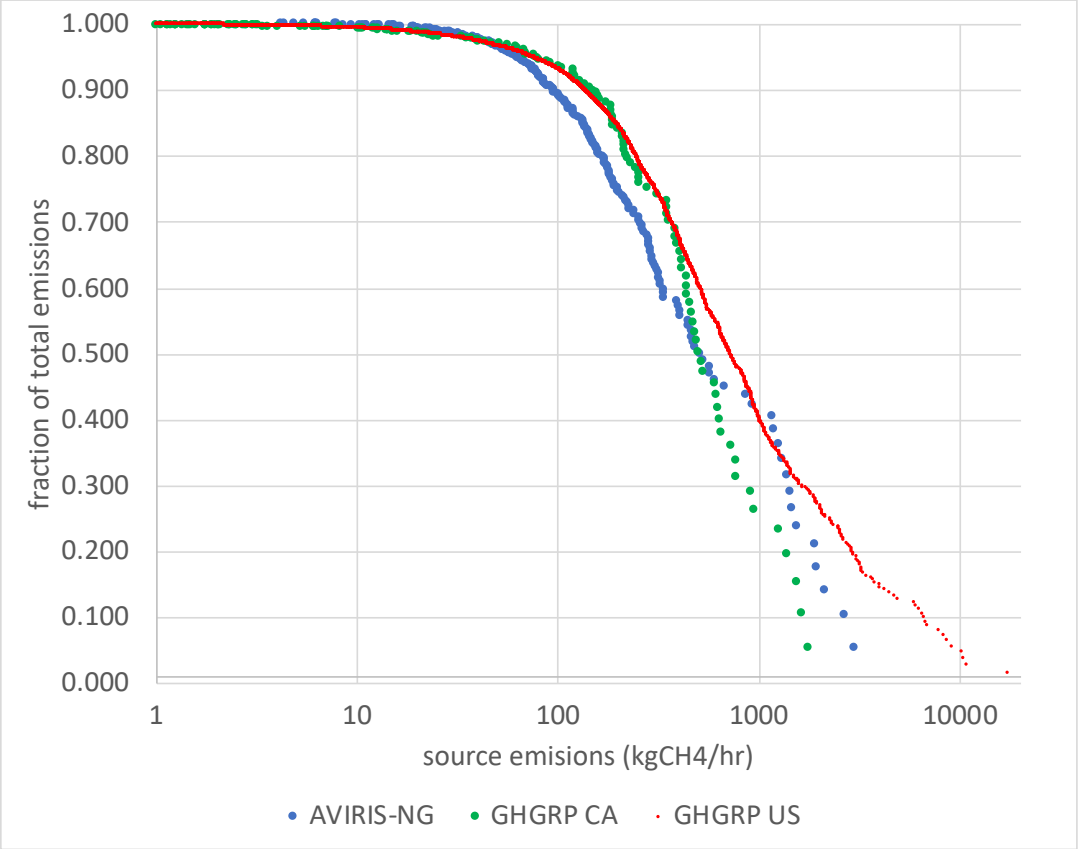

Figure S.14 Comparing 250 AVIRIS-NG facility scale emission with those with the EPA's 2017 GHGRP program for California (264 facilities) and the entire US (5800 facilities). The GHGRP curve for California indicates that 99% of the total originates from facilities emitting at least 10 kg h^{-1} whereas both the GHGRP curve for the total US and AVIRIS-NG curve indicate an equivalent completeness for facilities emitting at least 25 kg h^{-1}. This is significant considering that the AVIRIS-NG curve includes all point source sectors whereas key sectors such as manure management and oil and gas production typically do not participate in the GHGRP.

S4. Description of Supplementary Table files

PDF files containing:

- Supplementary Table S.4: Methane plume list. Source ID, latitude, longitude, detection date, detection time (UTC), source type, IPCC sector, IME/r (kg m^{-1}), $\sigma_{IME/r}$ (kg m^{-1}), U_{10} (m s^{-1}), σ_{U10}(m s^{-1}), Q_{plume} (kg h^{-1}), σ_Q (kg h^{-1}) [last two fields intentionally blank for those plumes lacking emission estimates due to quality control and filtering]
- Supplementary Table S.5: Methane source list. Source ID, latitude, longitude, source type, IPCC sector, number of overflights, persistence, confidence in persistence estimate, persistence adjusted average source emissions Q_{source} (kg h^{-1}), σ_Q (kg h^{-1})

Additional References for SI section

32. Carranza, V., Rafiq, T., Frausto-Vicencio, I., Hopkins, F. M., Verhulst, K. R., Rao, P., Duren, R. M., and Miller, C. E.: Vista-LA: Mapping methane-emitting infrastructure in the Los Angeles megacity, *Earth Syst. Sci. Data* **10**, 653-676 (2018).
33. Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K., Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing* **81** (5), 345-354 (2015).
34. Miller, C.M.F., P.L. Price, D. Meyer. Mass balance analyses of nutrients on California dairies to evaluate data quality for regulatory review. *Sci. Tot. Env.* **579**, 37–46 (2017).
35. California Integrated Water Quality System Regulated Facility Reports, https://ciwqs.waterboards.ca.gov/ciwqs/readOnly/CiwqsReportServlet?inCommand=reset&reportName=RegulatedFacility).
36. California Regional Water Quality Control Board (RWQCB), https://www.waterboards.ca.gov/publications_forms/publications/factsheets/docs/region_brds.pdf
37. Meyer, D., P.L. Price, H.A. Rossow, N. Silva-del-Rio, B.M. Karle, P.H. Robinson, E.J. DePeters, J.G. Fadel. Survey of dairy housing and manure management practices in California. *J. Dairy Sci.* **94**: 4744-4750 (2011).
38. USDA. Census of Agriculture; USDA-NASS: Washington, DC, 2012; quickstats.nass.usda.gov.
39. CDFA, California's Top 10 Milk Producing Counties Percent Share of California's Total Milk Production January-June 2017, https://www.cdfa.ca.gov/dairy/uploader/docs/DataChartsGraphs/Top_10_County_Milk_Production_Jan-Jun_2017.pdf
40. USDA. Survey of Agriculture; USDA-NASS: Washington, DC, 2015; quickstats.nass.usda.gov.
41. Thorpe, A.K., Frankenberg, C., Aubrey, A.D., Roberts, D.A., Nottrott, A.A., Rahn, T.A., Sauer, J.A., Dubey, M.K., Costigan, K.R., Arata, C. and Steffke, A.M., Mapping methane concentrations from a controlled release experiment using the next generation airborne visible/infrared imaging spectrometer (AVIRIS-NG). *Remote Sensing of Environment*, *179*, pp.104-115 (2016).
42. Benjamin, S.G., S.S. Weygandt, J.M. Brown, M.Hu, C. R. Alexander, T.G. Smirnova, J.B. Olson, E.P. James, D.C. Dowell, G.A. Grell, H. Lin, St.E. Peckham, T. L. Smith, W.R. Moninger, and J. S. Kenyon, A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Mon. Wea. Rev.* **144**, 1669-1694 (2016).
43. HRRR archive at the University of Utah, retrieved Aug 30, 2018, http://home.chpc.utah.edu/~u0553130/Brian_Blaylock/
44. NOAA HRRR archive, retrieved august 30, 2018, ftp://arlftp.arlhq.noaa.gov/pub/archives/hrrr/
45. Newman., J.F. and P.M. Klein, The Impacts of Atmospheric Stability on the Accuracy of Wind Speed Extrapolation Methods, *Resources* **3**, 81-105 (2014).
46. https://mesowest.utah.edu/

47. Varon, D., D. J. Jacob, J. McKeever, D. Jervis, B. O. A. Durak, Y. Xia, Y. Huang, Qantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes, *Atmos. Meas. Tech.* **11**, 5673–5686 (2018).
48. http://nomads.ncep.noaa.gov/cgi-bin/filter_rtma_ru.pl
49. Conley, S., Faloona, I., Mehrotra, S., Suard, M., Lenschow, D. H., Sweeney, C., et al., Application of Gauss's Theorem to quantify localized surface emissions from airborne measurements of wind and trace gases, *Atmos. Meas. Tech.* **10**, 3345-3358 (2017)
50. Timothy L. Vaughn, Clay S. Bell, Cody K. Pickering, Stefan Schwietzke, Garvin A. Heath, Gabrielle Pétron, Daniel J. Zimmerle, Russell C. Schnell, and Dag Nummedal, Temporal variability largely explains top-down/bottom-up difference in methane emission estimates from a natural gas production region, *PNAS* November 13, 2018 115 (46) 11712-11717
51. Jeffrey D. Wood, Robert J. Gordon, Claudia Wagner-Riddle, Biases in discrete CH_4 and N_2O sampling protocols associated with temporal variation of gas fluxes from manure storage systems, *Agricultural and Forest Meteorology* 171–172 (2013) 295–305
52. Taylor, Diane Margaret, Atmospheric Dispersion Modeling to Inform A Landfill Methane Emissions Measurement Method, Peer reviewed|Thesis/dissertation, 2017, https://escholarship.org/uc/item/24m0z749
53. Spokas, K., J. Bogner, M. Corcoran, S. Walker, From California dreaming to California data: Challenging historic models for landfill CH_4 emissions, *Elem Sci Anth* **3**, p.000051 (2015).
54. Conley, S. A., Faloona, I. C., Lenschow, D. H., Karion, A., Sweeney, C.: A Low-Cost System for Measuring Horizontal Winds from Single-Engine Aircraft, *J. Atmos. and Oceanic Tech.* **31**, 1312-1320 (2014).