Agriculture sector:
Enteric Fermentation and Manure Management Emissions from Individual Cattle Feedlots and Dairies

Aaron Davitt[1,5], Gabriela Volpato[1,5], Lee Gans[1,5], Xinnan Frank Cheng[1,2,5], Elliot Block[5], Krsna Raniga[1,5], Jackie Vandermel[5], Allan Mendoza[5], Dan McCrary[5], Akasha Sutherland[3], Reihaneh Rostami[3], Mitchell Smith[3], Brian Goodwin[3], Claire Pluard[4,7], and Sam Schiller[4,7]

1) WattTime, 2) Harvard University, 3) Synthetaic, 4) Carbon Yield, 5) Climate TRACE

1. Introduction

According to the Food and Agriculture Organization (FAO) data (FAOSTAT), beef and dairy milk production systems are the largest contributors of greenhouse gas (GHG) emissions in the livestock sector, representing more than 60% of emissions in the sector and 14.5% of all anthropogenic sources (FAO 2013). Beef and dairy sector emissions are driven by two sources. The primary source is enteric fermentation emissions which consists of methane (CH_4) gas produced in the digestive systems of ruminants and to a lesser extent non-ruminants. The secondary source is GHG emissions from manure management, producing both methane (CH_4) and nitrous oxide (N_2O) emissions via aerobic and anaerobic decomposition of livestock manure, including the microbially-driven processes of nitrification and denitrification (Waldrip et al., 2016; Waldrip et al., 2020). These emissions occur within manure storage facilities common to beef and dairy systems, as well as in-field where manure has been applied, or deposited by livestock.

The current de facto beef and dairy cattle emissions estimates are from the FAOSTAT and are based on country-level official, implicit, or estimated activity data and an indication of global cattle emissions in each country. While an indication at a global scale, such information is coarse and varies in data quality based on the country, and does not have specific facility attribution of emissions. Generally, facility level information is reported in academic studies, and, at times, the location information has been removed or kept vague (Harper et al., 2009; Costa et al., 2014; Zhu et al., 2014). However, while some jurisdictions have permitting databases that contain location data for livestock production data, the location of most livestock production facilities worldwide is unknown. As such, understanding facilities emissions and contribution to regional and global GHG emissions is a difficult process. Discrepancies between top-down and bottom-up assessments of methane emissions may be explained in part by these significant gaps in facility level assessments of livestock operations (Wolf et al., 2017).

In order to understand individual beef (hereafter, feedlots) and dairy emissions contributions, the Climate TRACE coalition has developed a method that employs an artificial intelligence tool called Rapid Automated Image Characterization (RAIC) and satellite imagery to identify feedlots and dairies in specific regions of interest in Russia, China, South Africa, Australia, Mexico, Brazil, Argentina, and the U.S.- California, Texas, Iowa, Michigan, and Nebraska where higher concentration of dairy and feedlots occur. Once identified, the feedlot area was defined and matched to permitting and reported cattle population databases, from various regions, that contain individual facility information to build a cattle population prediction model. The basis of this model is that the area size of a feedlot or dairy is related to the quantity of total cattle headcount confined within the area. Once each feedlot and dairy cattle headcount was estimated, the Intergovernmental Panel on Climate Change (IPCC) emissions factors (EFs) were applied based on the regional characteristics, temperature, cattle type, and manure management system to predict individual feedlot and dairy emissions for years 2015 to 2022. This approach has generated a first of its kind database of facility level emissions estimates for the cattle sector.

2. Materials and Methods

The approach utilized here primarily relies on the hypothesis that a feedlot or dairy area can be used as a predictor to estimate the total head of cattle, which can be converted to estimate enteric fermentation and manure management emissions. Previous studies have shown that number of animals per farm, or the total head of cattle, has a direct relationship to total GHG and non-GHG emissions (Harper et al., 2009; Vechi et al., 2022). This basis was applied to feedlots and dairies identified in these regions of interest (Figure 7),

1. Southwestern to Russia: ranging from Novooskolsky, Krai, and Arsky Districts.
2. Northeast China: mainly in Inner Mongolia, Shanxi, Hebei, Liaoning, and Henan.
3. Southwestern and northeastern South Africa: including Western Cape, Gauteng, and Limpopo.
4. Eastern Australia: Queensland and New South Wales.
5. Central Mexico: mainly in Jalisco and Queretaro.
6. West-central to central Brazil: State of Mato Grosso, Goiás, Minas Gerais, and São Paulo.
7. North-central Argentina: Cordoba, Buenos Aires, and Santa Fe Provinces.
8. U.S.: California, Texas, Iowa, Michigan, and Nebraska.

These geographies were selected due to their high concentration of dairy and beef feedlots in their respective country. Not all the feedlots and dairies were identified in these regions, however the ones that were identified do represent some of the largest cattle operations in these regions.

A combination of permitted/reported total head of cattle by type, artificial intelligence, satellite imagery, modeling, and emission factors were used to estimate emissions for 2015 to 2022 for each identified feedlot and dairy in each region of interest. The sections below provide a description of each dataset employed in this study.

2.1 Datasets employed

To create the Climate TRACE feedlot and dairy emissions dataset, the following data sources and information were accessed and employed for feedlot and dairy identification, model development, and to attribute specific emission factors by cattle type at individual locations.

2.1.1 Remote sensing datasets

The following satellite imagery datasets were ingested in RAIC (see section 2.2) to identify feedlots and dairies:

To identify feedlots and dairies within the U.S., The National Agricultural Imagery Program (NAIP) aerial imagery distributed by The U.S. Geological Survey (USGS). NAIP acquires aerial imagery in the red, blue, green and near infrared wavelengths at a 1m spatial resolution, or finer, as part of an agricultural census conducted in the U.S. Imagery is available every three years beginning in 2009. NAIP imagery acquired in 2020 to 2022 (year image varying by the state) were used to identify feedlots and dairies in California, Texas, Iowa, Michigan, and Nebraska with the RAIC tool (described in section 2.2 and 2.3.1). More information on NAIP imagery can be found at the USGS NAIP imagery program page (U.S. Geological Survey, 2018).

To identify feedlots and dairies outside the U.S., visual basemaps imagery created by Planet Lab's PlanetScope satellite constellation was accessed for the RAIC tool (Planet, 2022). Each basemap was generated from optimal PlanetScope imagery comprising blue, green, red visual imagery. More information on Planet basemaps can be found on their website (https://developers.planet.com/docs/data/visual-basemaps/). For the purposes here, mosaiced basemap imagery from 2020 and 2022 were used. To account for seasonality that can lead to increased cloud cover and hazy conditions, the summer months' basemap were accessed and used with RAIC (described in section 2.2 and 2.3.1). Planet imagery was selected for regions outside the U.S. due to the lack of freely available global satellite imagery data with a spatial resolution less than 6m or finer.

2.1.2 Permit and reported feedlot and dairy data

To assess the feedlot and dairy area size relationship to total cattle at a facility, total cattle head count was accessed from various sources, shown in Table 1. In addition to total cattle head count, some of these databases reported additional information at specific locations- cattle type, facility area, manure management types, and cattle subtypes at individual locations. Some of these databases were filtered and cleaned to contain only beef and dairy cattle information. This required using the Standard Industrial Classification (SIC) and North American Industrial

Classification System (NAICS) codes, or by reported animal type (i.e., the number of milking cows at a location), or keywords in the ownership name - “feedlots”, “calves”, “beef” etc.

Table 1 Permitting and reporting cattle databases with information related to model development and emissions estimates listed in the columns. All datasets, except Fan *et al.* (2018), have latitude and longitude information. Links to sources are in the references section.

Country/ Region	Source	Cattle headcount	Cattle type	Facility area	Manure management type	Other information	Accessed
Australia	The Farm Transparency Map	no	no	no	no	Cattle/CAFO subtypes	Sept. 2023
China	Fan *et al.* (2018)	yes	yes	yes	no	Cattle/CAFO subtypes	Feb. 2023
South Africa, Western Cape	Department of Environmental Affairs and Development Planning Cattle Census	yes	no	no	no	Ownership information; Cattle/CAFO subtypes	Jan. 2023
U.S., California	California Integrated Water Quality System Project (CIWQS)	yes	yes	no	no	Ownership information; SIC/NAICS; Cattle/CAFO subtypes; Year active	April 2023
U.S., California	Sources of Methane Emissions (Vista-CA), State of California, U.S. (modified CIWQS data; Hopkins et al. 2019)	no	no	no	no	Ownership information	April 2023
U.S., Iowa	Iowa Department of Natural Resources	yes	yes	no	yes	Ownership information	July 2023
U.S., Michigan	Department of Environment, Great Lakes, and Energy	yes	yes	yes	no	Ownership information; Cattle/CAFO subtypes	June 2023
U.S., Nebraska	Nebraska Dept. of Agriculture	yes	yes	no	no	Ownership information	May 2023

U.S., Texas	[Texas Commission on Environmental Quality]	yes	yes	no	yes	Ownership information; Cattle/CAFO subtypes; Year active	Feb. 2023

2.1.3 Geosearch tool to identify cattle operation types

For regions and countries that do not have publicly available cattle databases, Climate TRACE developed a tool called "Geosearch" to query other search engines and datasets. Geosearch ingests identified cattle facilities and then queries multiple data sources based on the latitude and longitude information, and presents results in a unified user interface. The core data integrations included are shown in Table 2.

Table 2 Geosearch core data integrations to enable cattle operation identification.

Source	Information
Google Maps	Imagery (present-day); Location Metadata
Planet Labs	Imagery (historical) to identify when a facility was active
Google Places	Location metadata for ownership information and total head of cattle (if available)
Bing Search	Web pages (based on Location name, coordinates) for ownership information and total head of cattle (if available)
Google Translate	Language translation for location metadata and web pages, such as Chinese and Russian web pages

Using a multimodal, single-search interface allowed comparing data sources side-by-side, as well as joining data from one source to another – e.g. combining a list of Google Places with a sub-list of Bing Search results for each place – without leaving the interface. Human annotators used the tool to search the web for additional data, and visually inspected historical remote sensing imagery, in order to manually annotate the identified cattle operations by feedlot or dairy and additional metadata - total head of cattle, ownership, the year facility was built (year active), and the presence of pollution control technologies such as anaerobic digesters.

2.1.4 Total cattle capacity utilization at feedlots and dairies (capacity utilization factor)

Feedlots and dairies are designed to hold a certain maximum total number of cattle. Generally, dairies maximize and maintain the number of dairy cows milked for production, with large dairies typically operating more efficiently in terms of labor and resources used. However, feedlots have a number of factors that can limit the number of cattle occupying the feedlot space. This includes market demands, cost of feed, and the weight of the animal prior to feedlot placement (Turcios 2022). The last factor - weight of the animal prior to feedlot placement -

reflects the fact that feedlots are "through systems" where beef cattle only spend a fraction of their time on feedlots where they are fattened for slaughter (Hayek and Garret 2018; USDA 2023). The fattening process can be influenced by market demands and the price of feed. As such, these factors influence the number of cattle on a feedlot, or capacity utilization, generally resulting in less than 100% of the space utilized. To capture the difference in capacity utilization at dairies and feedlots yearly, the following approaches were employed. First, it was assumed that the total head of cattle at dairies was maximized for milk production. Therefore, the total dairy cattle population capacity utilization was set to "1" in the model input to represent the dairy operating at full capacity, or 100% capacity. For feedlots, literature research was performed to identify capacity utilization factors for different countries (Table 3). Feedlot capacity utilization values range between 0%, no beef cattle present, to 100% utilization. These capacity utilization factors were applied to feedlots in their respective regions. For regions where capacity utilization factors were unavailable (i.e., Russia and Argentina) the mean value from all identified sources in Table 3 was applied to these regions for each year.

Table 3 Feedlot capacity factors by region.

Country/Region	Source	Capacity factor
Australia	Meat & Livestock Australia Limited Lot Briefing Reports (Various years)	Each year's capacity factor identified in the reports
U.S.	U.S. Dept. of Agriculture, Iowa State Univ. (2018)	For each year, the capacity factor was based on the total number of cattle on feed divided by total feedlot capacity in the U.S. per year using USDA data
Mexico	Peel et al. (2011)	A single mean capacity factor was created from a range and applied to all years
China	Waldron et al. (2015)	A single mean capacity factor was created from a range and applied to all years
South Africa	Pienaar et al. (2019)	A single mean capacity factor was created from a range and applied to all years
Brazil	JBS Foods (Pereira 2023) and Minerva Foods	Slaughter capacity utilization rate was used as a proxy for each year and applied based on feedlot proximity to slaughterhouse (Skidmore et al., 2022)

2.1.5 IPCC emission factors

EFs from The IPCC "Chapter 10: Emissions from Livestock and Manure Management" and "Chapter 11: N_2O Emissions from Managed Soils, and CO_2 Emissions from Lime and Urea Application" (IPCC 2006a; IPCC 2006b) were used to translate cattle populations into emissions estimates. Cattle emissions are produced by enteric fermentation, producing CH_4, and manure management, producing both CH_4 and N_2O. To estimate CH_4 and N_2O emissions, the IPCC Tier

1 approach was applied with a Tier 2 approach included for Indirect N_2O emissions due to leaching from manure management only. For dairies and feedlots that lacked detailed manure management systems, assumptions were applied, discussed further in section 2.3.2.

Default IPCC EFs based on region, temperature, and manure management systems were applied to the total head of cattle at individual feedlots and dairies. Additionally, the uncertainty estimates, were expressed as a percentage above or below the mean estimate (i.e. +/-XX%), or as an interval with an upper and lower bound of values. For example, Table 10.14 “Manure management methane emission factors by temperature” states these EFs have an uncertainty of ±30 % (IPCC 2006a). For the uncertainty range for leaching from manure management systems, a lower, 1%, and upper, 20%, range was used from the typical reported range of 1 - 20% (IPCC 2006a). See Table S2 for more information on the IPCC uncertainty estimates.

2.1.6 Temperature data

Appropriate application of IPCC EFs to estimate manure methane emissions required temperature data for each modeled facility. The average annual temperature for each individual feedlot and dairy was produced by ERA5-Land Daily Aggregated from the European Centre for Medium-Range Weather Forecasts (ECMWF) Climate Reanalysis via Google Earth Engine (Muñoz Sabater, J., 2019). The specific image collection accessed was “ECMWF/ERA5_LAND/DAILY_RAW”, which provided temperature data up to 2022 data, needed for this work (ERA5-Land Daily Aggregated data was accessed February 2023). The “temperature_2m” band was used to derive the average annual temperature for years 2015 to 2022 for each feedlot and dairy. This value was used to determine the IPCC “Manure management methane emission factors by temperature” in Table 10.14 (IPCC 2006a). For feedlots and dairies with no temperature data, the closest feedlot or dairy with temperature data was used.

2.2 Rapid Automatic Image Categorization (RAIC) tool

Identification of feedlot and dairy facilities was performed in partnership with Synthetaic, a specialist in geospatial applications of machine learning tools including the Rapid Automated Image Characterization (RAIC) approach (https://www.synthetaic.com/). RAIC is a proprietary artificial intelligence tool developed by Synthetaic. By providing RAIC with raw, unlabeled imagery, this tool can be applied to large volumes of data to automatically find objects of interest. Additionally, using a human nudge tool, users can further refine the AI to better identify areas that contain a user-defined object of interest. For the work here, using a set of images of pre-validated cattle operations facilities in each region, RAIC searched millions of square kilometers of remote sensing imagery and identified operations with similar features relative to the seed data, such as color, shape, and orientation. Once identified, a human-review process followed to validate the results. Figure 1 provides an overview of the RAIC process.

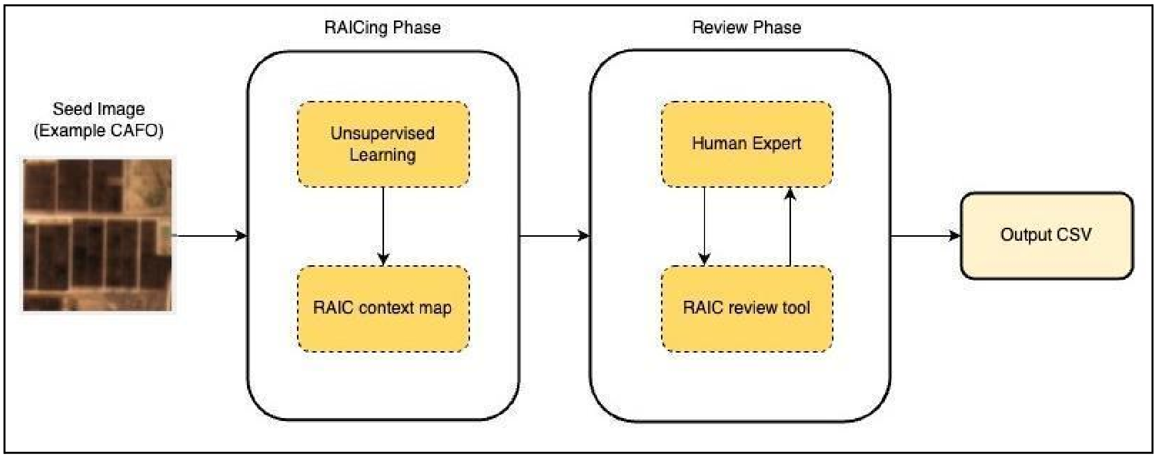

Figure 1 An example of the RAIC process to identify feedlots and dairies. Further description in text below.

RAIC detection results were presented as a list of entries in a comma-separated values (CSV) file, with each entry referring to a specific feedlot or dairy tile found in RAIC. In many instances, a whole feedlot or dairy can cover a vast array of tiles, depending on the resolution of data ingested (Figure 2). For example, the same feedlot in NAIP data would be fewer tiles in PlanetScope data, given PlanetScope is ~3m and NAIP is ~1m. (Figure 5A and B). Following the export of the CSV values, there was a need to consolidate those results into groups, where each group represented a whole cattle operation. For that reason, Synthetaic created the Nearest Neighbors algorithm (not to be confused with the supervised K-NN learning algorithm). It was a process where all adjacent tiles were consolidated into a singular group. Following this, groups whose centroids were less than a minimum distance away from each other were merged to form a singular group of tiles representing the entire feedlot or dairy.

Each tile was assigned a unique tile ID as well as latitude and longitude coordinates that represent that tile's centroid. After a unique ID was generated, and the distance from each tile, *T*, was calculated to every other tile in the results CSV file. Once this distance was calculated, and compared to a minimum distance threshold, it was determined which tiles were adjacent neighbors. The immediate neighbors were grouped and assigned a unique ID.

At this point (the "Merging" stage), groups were checked to determine if they are too close to another group where they might also be associated. For each group, we calculated an average latitude and longitude based on its member tiles and checked with the Haversine distance, accounting for the shape of the Earth. If two or more groups were closer than ~1km, they were merged. The new merged group assumes a singular unique Group ID. The distance threshold of ~1km was determined to be a sufficient distance based on internal reviews of the data. There may be some cases where a few groups will be farther away from each other (>1km) yet still belong to the same cattle operation.

The final resultant file from this "nearest neighbors" process was a feature list of detections in a CSV format, with each tile having an additional Group ID column specifying which group the detection belongs to. Additionally, the Group Size was calculated for each group representing the number of tiles belonging to that Group ID. This can be used, for example, to detect the number of large cattle operations, e.g. exceeding 10 tiles (Figure 2).

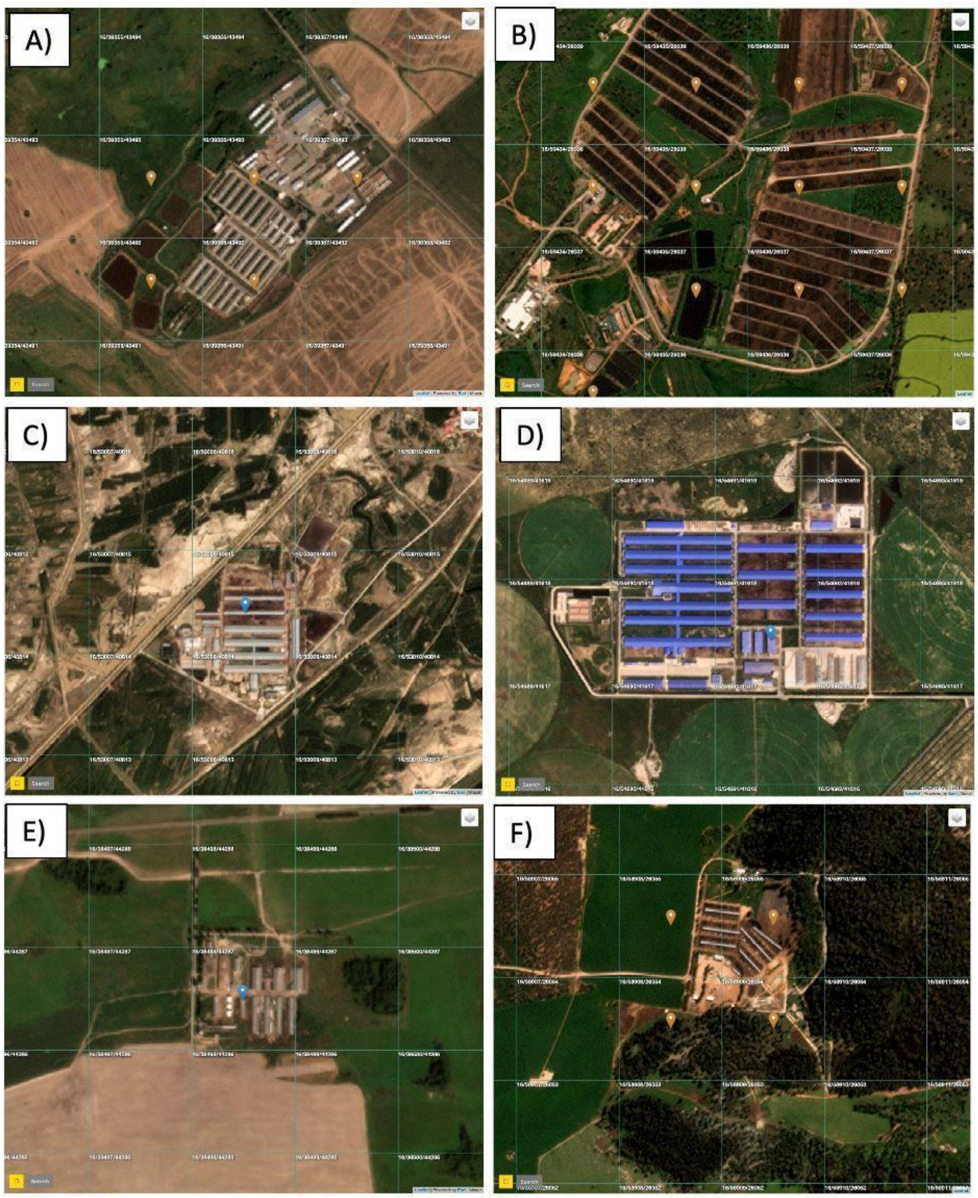

Figure 2 Examples of the diverse cattle operations in Australia, China, and Russia. A) A small (~7 tiles) and B) large (~17 tiles) feedlots in Australia, respectively. C) A small (~4 tiles) feedlot relative and D) large (~9 tiles) dairy in China. E) A small (~2 tiles) and F) large (~4 tiles) feedlot

in Russia. Cattle operation features that extend over more than one tile were grouped together to represent a single feedlot or dairy.

2.3 Methods

Figure 3 gives an overview of the model development and deployment which used feedlot and dairy area to predict total head of cattle which was converted to enteric fermentation and manure management emissions using IPCC EFs. Each step is described in further detail below.

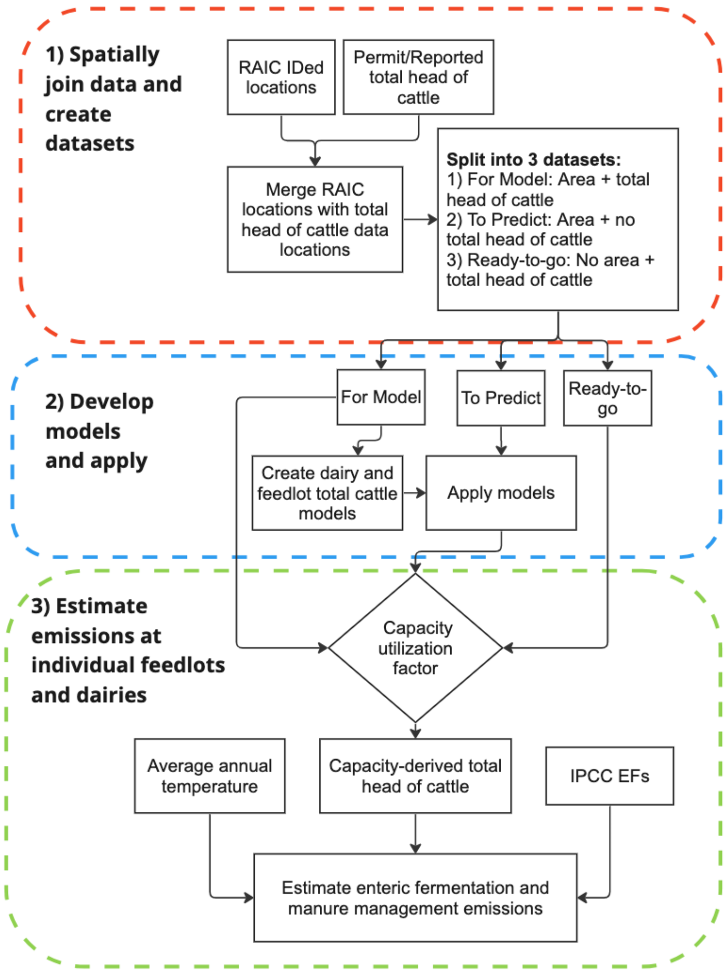

Figure 3 Flowchart summarizing each section's inputs and outputs to estimate total head of cattle and GHG emissions. 1) *Spatially join data and create datasets*: RAIC identifies individual feedlots and dairies, and boundaries were drawn to estimate the area in hectares and exported to a shapefile. This shapefile was spatially merged with datasets containing the total head of cattle. This created three datasets, described further in text. *2) Develop models and apply*: feedlot and dairy models were developed to estimate total head. 3) *Estimate emissions at individual feedlots and dairies*: each feedlot and dairy had a capacity utilization factor applied to scale their estimated total head of cattle. Annual mean temperatures and EFs were applied to estimate emissions.

2.3.1 RAIC utilization

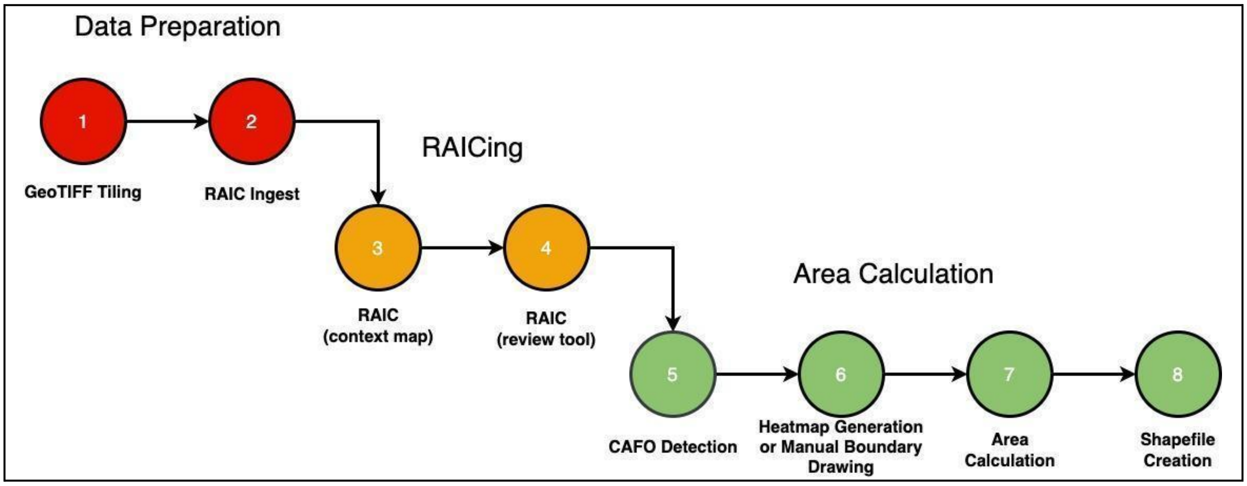

Figure 4 RAIC utilization to identify feedlots and dairies. Description is included in the text below.

The RAIC toolset was utilized to identify feedlots and dairies in regions of Russia, China, South Africa, Australia, Mexico, Brazil, Argentina, and the U.S.- California, Texas, Iowa, Michigan, and Nebraska, visualized in Figure 4. This involved the following steps:

1) *Data preparation* - NAIP or PlanetScope basemap imagery was downloaded and tiled (mosaiced) for the region of interest. In each region, either permit/reported data or research was conducted. This included using Google Maps and typing in keywords (beef, dairy, cattle etc.) in the region's native language. If a cattle operation was identified, the CAFOs latitude and longitude coordinates were input into RAIC - in order to produce a seed point. Using this initial seed point, the RAIC process was continued to identify similar visualizations to generate more seed points that can be used. Some countries - China and Russia - required a number of significant seed points because of the diverse cattle operations identified (Figure 2).

2) *RAICing* - Utilizing the seed points generated from Step 1, these were provided to RAIC to identify features related to a feedlot and dairy in the NAIP or PlanetScope basemap imagery. Then the RAIC tool proposed remote sensing tiles which had a high probability of containing feedlots and dairies in other NAIP and PlanetScope imagery. Human reviewers, within the RAIC tool, filtered what was and was not a feedlot or dairy to further refine the proposed tiles.

3) *Area calculation* - Once the RAICing process was completed, the feedlot or dairy boundary was generated using GIS polygon tools (Figure 5). Polygons could then be translated into facility area estimates, in hectares (ha), that can be used as a predictor to

estimate the total cattle headcount at an individual location. The boundary drawing was performed two ways based on specific regions.

- For California, Texas, and Argentina- these three regions were part of the initial study for this sector. In these regions the RAIC tool drew a convex hull heat map to calculate the boundary area (Figure 5A and B). This was done by calculating the land mass inside each convex hull to generate an area estimate for each identified cattle operation. As a result, some operations had their boundaries over and underestimated the size of the cattle operation, leading to larger or small area sizes. This is due to the challenges of separating the feedlot or dairy common features from physical features of the surrounding land. To improve the representation of a cattle operation's area, manual drawing was performed to better represent a feedlot or dairy. Only a subset of California and Texas was completed, not Argentina, which also included the identification and drawing of manure ponds at each operation. This provided an indication of manure management systems employed at a location.

- For Mexico, Brazil, South Africa, Australia, China, and Russia- these regions had their feedlots and dairies polygons drawn manually to represent their area size for the 200 largest facilities, based on tiling groupings described in section 2.2. Additionally, the manure retention ponds (visually identified adjacent to the facility and not separated by large swaths of land or roads) were manually drawn at these operations to provide an indication of manure management systems employed.

Once drawn boundaries were completed, each feedlot and dairy identified was exported to a shapefile and the area in ha derived.

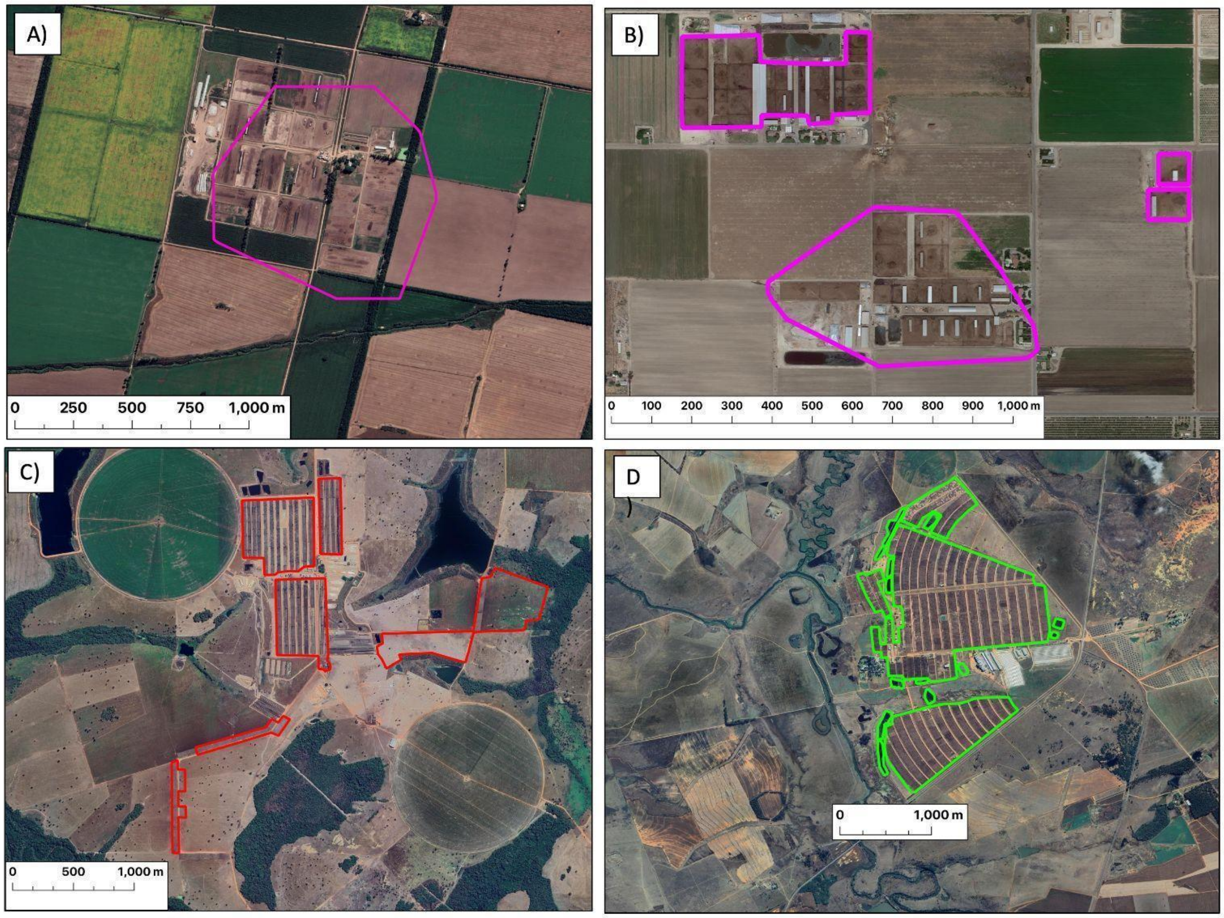

Figure 5 Examples of boundary areas shapefile creation in A) Argentina, B) California, C) Brazil, and D) South Africa. Argentina and California area boundaries were drawn using the "convex hull" process, leading to over and underestimating of area size. An example of a fixed boundary is shown in the California image (boundary top left). Brazil and South Africa area boundaries were drawn manually.

2.3.2 *Process permit/reported data and merge datasets*

To generate a total head of cattle model for feedlots and dairies required an area (predictor) and reported total head of cattle by type (predictand). At the time of this research, no known dataset was available that contained both. Therefore, Climate TRACE created its own dataset to estimate enteric fermentation and manure management emissions. This section refers to the *1) Spatially join and create datasets* shown in Figure 3.

The RAIC generated shapefiles were spatially merged with permit/reported datasets to create a curated feedlot and dairy database containing RAIC-derived area estimates and information from datasets in Table 1. All the databases in Table 1 were processed and cleaned to ensure only cattle-related information was contained. Additionally, further processing was performed including the removal of duplicate locations and creation of a cattle type tag (beef or dairy), based on publicly available cattle operation information. To ensure there was a large enough sample size for model training, the Geosearch tool was employed to manually identify cattle operations by type outside the U.S. Some facilities required using Google maps and street view and/or searching for the business address to provide information needed for modeling. Each permit/reported dataset had nuances that required region specific processing for use. If no cattle type information was available to assign to a cattle operation, these locations were automatically assigned as feedlots with beef cattle.

Once processing was completed, the RAIC identified feedlots and dairies were spatially joined to permit/reported data that created the following datasets:

A) *For Model* - this dataset contained spatially joined feedlot and dairy areas and total head of cattle for each location. This created 'one-to-one' matches used to develop dairy and beef (feedlot) models to predict total head of cattle for locations in the "*To Predict*" dataset. This dataset contained 710 cattle operations for model training.
B) *To Predict* - dataset contains area estimates only and no total head of cattle information but has cattle operation type. This dataset contained 2,056 cattle operations to model.
C) *Ready-to-go* - dataset contains no area but has total head of cattle and cattle operation type. This dataset has no spatially matched RAIC locations but contains the total head of cattle by type needed for emission estimates. Note, for the feedlot or dairy area size for these locations the value was set to "1". This dataset contained 4,592 cattle operations ready for emission estimates.

In all the datasets described above, the identified manure management systems were translated into IPCC equivalents to estimate N_2O emissions, shown in Table S1. In the majority of the systems identified, the IPCC description was used to categorize identified manure management systems but some assumptions had to be made. For example, it was assumed if a system is liquid-based, it was assigned to "liquid slurry"; anything related to aerobic treatment was assigned to "aerobic treatment". Additionally, any location with an anaerobic system, which IPCC assigns a 0 EF, was assigned as liquid/slurry emission since previous work indicates that this type of system may underestimate emissions (Owen and Silver, 2011; Petersen 2018) If the manure management system was unknown, it was assumed some manure handling was performed and assigned as "Drylot". For RAICed feedlots and dairies with no reported manure management systems, a subset had manure pond boundaries drawn, which can include various liquid storage systems. As a default, these locations assumed "liquid/slurry" systems.

In total, 7,358 cattle operations were identified for emission estimates (Figure 6). Of the total, 4,569 were feedlots and 2,789 were dairies. Due to the PlanetScope tiles extending outside of the region of interest, other countries' cattle operations were included - Botswana, Kazakhstan, and Mongolia. The breakdown of cattle operations per country is shown in Table 4.

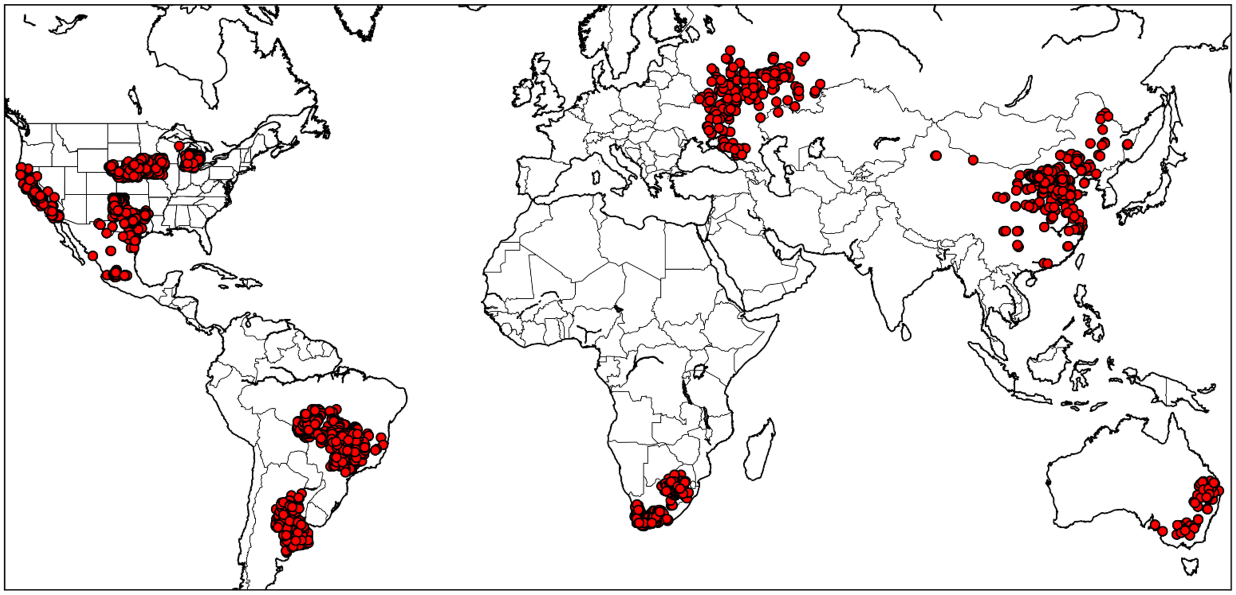

Figure 6 Regions of interest where feedlots and dairies were identified.

Table 4 Number of cattle operations per country.

Country	Number of Cattle Operations	Country	Number of Cattle Operations
Argentina	378	*Mexico*	90
Australia	90	*Mongolia*	1
Botswana	1	*Russia*	186
Brazil	403	*South Africa*	2,750
China	455	*U.S.*	2,990
Kazakhstan	1	*Ukraine*	13

While this data cleaning and joining was performed for all regions to minimize error and perform quality control, mislabeled and misidentified cattle operations may exist and we welcome users to report these operations if identified.

2.4 Modeling cattle populations at each facility

This section refers to the *2) Develop models and apply* in Figure 3.

2.4.1 Feedlot and dairy model development

The model development relies on the relationship between feedlot and dairy area relationship to the total head of cattle at a location. Using the "For Model Training" dataset described in section 2.3.2, our analysis found that the feedlot or dairy area and the total head of cattle have a direct relationship - as the area size increases, the total head of cattle increases. See the Results section and Figures 7 and 8 for more information. Using this relationship, a total of four models were developed based on based on specific regions:

- *Feedlot U.S. Southwest Model* - this model was based on the relationship between Texas and California feedlot areas to the total head of beef cattle at each operation.
- *Dairy U.S. Southwest Model* - this model was based on the relationship between Texas and California dairy areas to the total head of dairy cattle at each operation.
- *Dairy U.S. Midwest Model* - this model was based on the relationship between Michigan dairy areas to the total head of dairy cattle at each operation.
- *Dairy China Model* - this model was based on the relationship between China dairy areas to the total head of dairy cattle at each operation.

All four models are linear regression models:

$$Total\ Head\ of\ Cattle_{i,r} = (m_{t,r} * Area_i) + B_{t,r}\ (Eq. 1)$$

Where, the *Total Head of Cattle* at an individual feedlot or dairy, "*i*", in a region of interest (i.e, China, U.S.), "*r*", is the function of:
"*Area*": the size of the feedlot or dairy, in hectares.
"m": the slope of the regression line based on how the total head of cattle changes relative to a change in feedlot or dairy area size. This slope is region-specific to cattle operations type, "*t*", feedlot or dairy.
"B": the region-specific constant based on cattle operations type.

2.4.2 Modeled estimates

Feedlots and dairies with no total head of cattle data, the "To Predict" dataset, were modeled using Eq. 1. Specific models were applied to specific regions to estimate total head of cattle based on feedlot or dairy area size:

- *Feedlot U.S. Southwest Model* - was applied to all identified feedlots in all regions.
  This model was applied to U.S. and Mexico feedlots, because of The North American Free Trade Agreement (NAFTA) impact on integrating the North American market (Zahniser 2015; OSU 2018). Additionally, this model was applied to all feedlots globally since all have similar features - large open pens and dirt lots.
- *Dairy U.S. Southwest Model* - this model was applied to dairies identified in Texas, Nebraska, and California, U.S., and Mexico, Brazil, and South Africa.
  This model was applied to dairies in these regions due to similar features seen in all U.S. states and Mexico - milking parlors and numerous confined facilities within a dairy. This was applied to Mexican dairies since they are located near the U.S. as NAFTA has had an

impact on increased milk production (Zahniser 2015; OSU 2018). For Brazil and South Africa, this model was applied due to the lack of identified dairy facilities in each country.

- *Dairy U.S. Midwest Model* - this model was applied to dairies identified in Iowa and Michigan, U.S.
- *Dairy China Model* - this model was applied to dairies identified in China only.

In some cases, a filled in value was applied to feedlots and dairies with an area size less than 10 ha. These smaller cattle operations that had their total head of cattle modeled resulted in negative values, depending on the region. This was due to the Climate TRACE goal to identify the largest feedlots and dairies, thereby the largest emitters, in each region. As a result, the models were biased towards larger cattle operations. Any cattle operation with a negative value had a mean value applied for their total head of cattle estimates. This mean fill-in value was generated from feedlots in South Africa, Brazil, California, and Texas with feedlots that have cattle and are less than 10 ha in area size and applied to all cattle operations with a negative value.

2.4.2.1 Permit/reported estimates

Feedlots and dairies with reported total head of cattle were used as is and no modeling was performed for these datasets. This includes the "For Model" and the "Ready-to-go" datasets.

2.4.3 Cattle population utilization factor and emissions estimates

This section refers to the *3) Estimate emissions at individual feedlots and dairies* in Figure 3.

Feedlot total head of cattle can be influenced by factors that can reduce the total cattle within a feedlot (section 2.1.4). As such, a feedlot capacity utilization factor was applied based on regional information, shown in Table 3. This adjusted the total head of cattle estimates in Eq.1 to the following:

$$Actual\ Total\ Head\ of\ Cattle_{i,r,yr} = Total\ Head\ of\ Cattle_{i,r} * \ cuf_{t,r,yr}\ (Eq.\ 2)$$

Where, the *Actual Total Head of Cattle* for a feedlot or dairy in a region is the product of the *Total Head of Cattle* from Eq.1 and the *cuf*, the capacity utilization factor, for a feedlot or dairy in a region for a specific year, *yr*. The *cuf* for all dairies, in all regions for all years, was set to 1, indicating the dairies were at full utilization. Feedlots' *cuf* varied by region and year but values ranged between 0<x<1 based on values derived from sources in Table 3.

Once the *Actual Total Head of Cattle* per dairy and feedlot was generated, IPCC equations and default regional EFs were used to estimate enteric fermentation and manure management emissions.

2.4.3.1 Enteric fermentation emissions

For enteric fermentation, Table 10.11 EFs were applied (IPCC 2006a). The “Other cattle” EF was applied to feedlots, assuming all were beef cattle, and equation 10.19 was used to estimate the total feedlot CH4 emissions (IPCC 2006a). For dairies, two approaches were taken since two cattle types were reported at specific locations - total milking cattle and total non-milking cattle. First, if the facility reported only a total milking cattle population, then the “Dairy” EF was applied. Second, if a facility reported a total milking cattle and total non-milking cattle populations, then the “Dairy” EF was applied to the total milking cattle population and the “Other cattle” EF was applied to the total non-milking cattle population. Both enteric fermentation emission estimates were estimated using equation 10.19 and summed to estimate total dairy CH4 emissions (IPCC 2006a).

2.4.3.2 Manure management emissions

For manure management emissions, two pathways were used to estimate CH_4 and N_2O emissions. CH_4 emissions from manure handling are affected by temperature. The 2015 to 2022 ERA5 annual mean temperature data generated for each feedlot and dairy was used to find the nearest temperature-based EF in Table 10.14 (IPCC 2006a). The “Other Cattle” EF was applied to feedlots and equation 10.22 applied to estimate the total CH_4 emissions value. As described in section 2.4.3.1, if a dairy had only total milking cattle, then the “Dairy Cows” EF was applied, or if total non-milking cattle populations was known, then the “Dairy Cows” and “Other Cattle” EFs were applied and equation 10.2 was summed to estimate total CH_4 emissions (IPCC 2006a).

N_2O emissions from manure management can be produced directly and indirectly - volatilization and leaching - from the systems employed at a feedlot or dairy (IPCC 2006a). To estimate direct and indirect emissions, first, using each feedlot or dairy’s location, the default regional dairy or other cattle nitrogen excretion rates (Nex) from IPCC Table 10.19 was applied to estimate total Nex (IPCC 2006a). Then for each individual feedlot or dairy, the manure management system was either known, based on information from Table XX, or estimated. If estimated, the following assumptions in manure management systems were applied:

- *For unknown manure management systems in the U.S.*: dairies were assumed to have one liquid slurry and one aerobic treatment system. At unknown feedlots, we assumed one liquid slurry system. These assumptions were based on the most frequently reported manure management systems identified in the U.S. permit/reported data in Table XX.
- *For unknown manure management systems for the rest of the world (ROW)*: If the dairy was in China (including the one identified in Mongolia) or Mexico, then it was assumed a liquid slurry system based on Table 10A-4 (IPCC 2006a) . For the rest of the dairy ROWs (South Africa, Australia, Brazil, and Argentina) and ROW feedlots, all were assumed to have one dry lot manure management system since the manure must be handled and stored in some form.

To estimate direct N_2O emissions, IPCC equation 10.25 was used; indirect emissions due to volatilization used equations 10.26 and 10.27; and indirect emissions due to leaching used equations 10.28 and 10.29. The manure managements system(s) at each feedlot or dairy was used to select the manure management EFs from Table 10.21 and 11.3, along with nitrogen loss due to the volatilization of NH_3 and NO_x from the manure management system ($Frac_{GasMS}$; Table 10.22) and the percent of managed manure nitrogen losses for livestock category T due to runoff and leaching ($Frac_{LeachMS}$) were applied to the equation where required (IPCC 2006a and 2006b). The $Frac_{LeachMS}$ value was set to a fixed value of 10.5% based on the reported typical range of 1-20% (IPCC 2006a). Once each direct and indirect emissions estimate was derived, then each one was multiplied by the number of cattle.

To represent on-site manure management emissions, modifications were made to the IPCC approach to reflect more detailed information available. First, the "fraction of total annual nitrogen excretion for each livestock species/category T that is managed in manure management system S in the country, dimensionless" or $MS_{(T,S)}$ was set to 1 since the emissions from the manure management itself was estimated. Second, there were feedlots with a liquid/slurry system but no IPCC "Other Cattle" $Frac_{GasMS}$ for liquid/slurry. Instead, any feedlot with a liquid/slurry system used the "Dairy Cow" $Frac_{GasMS}$ for liquid/slurry. This application also extends to any liquid-based manure management system identified and if a dairy operation had reported non-milking cattle. Lastly, if a feedlot or dairy had multiple manure management systems of the same type and/or more than one manure management systems, then each direct and indirect emissions estimate was multiplied by the number of respective manure management systems at that location, and then summed to generate a total N_2O emissions.

2.5 Reporting of emissions on Climate TRACE

In total, 7,358 cattle operations had their enteric fermentation and manure management emissions estimated, of which 4,569 were feedlots and 2,789 were dairies. For South Africa, we had to aggregate emissions to the province level to maintain ownership anonymity. As a result, this produced 4,609 individual feedlots and dairies' emissions globally and 31 South African provinces with cattle operation's emissions estimates.

On the Climate TRACE website, enteric fermentation and manure management emissions were reported separately since the United Nations Framework Convention on Climate Change (UNFCCC) reports emissions as Agriculture 3.A.1.a Enteric Fermentation - Cattle and 3.B.1.a Manure Management - Cattle. Default emissions values were reported as CH_4, N_2O, and CO_2 equivalent 20- and 100-year global warming potential (CO_2e 20yr and 100yr GWP) on the Climate TRACE website. Additionally, each feedlot and dairy's enteric fermentation and manure management emissions had their uncertainty estimates produced. The uncertainty values used are described in section 2.1.5.

Lastly, we included confidence and uncertainty estimates for different data fields. Confidence estimates were provided on a 5-point scale: very low, low, medium, high, and very high. This is to indicate how certain we are of a data field used to generate emissions. Additionally, uncertainty estimates were provided, based on each model's standard deviation, and IPCC uncertainty percentages and ranges. The Supplementary metadata section provides an overall description of the emissions data created in Table S1 and the confidence and uncertainty emission estimates in Table S2.

2.6 Verification of approach

The following verifications were performed to evaluate the approach to estimate total head of cattle at feedlots and dairies. To evaluate the relationship between cattle operation's area and reported total head of cattle a Spearman's rank correlation coefficient (Rs) was computed to assess the strength of the relationships between the two. Additionally, linear regression plots display the relationship between area to total head of cattle including the mean square error (MSE) and the goodness-of-fit measure (R^2), root mean square error (RMSE), and mean absolute error (MAE).

3. Results

3.1 The relationship between feedlot and dairy area to total head of cattle

The model development relies on the relationship between cattle operations area to the total head of cattle at each location. Using the "For Model Training" dataset described in section 2.3.3, our analysis found that the feedlot or dairy area and the total head of cattle have a direct relationship. As the area size increases, the total head of cattle increases. For feedlots, the overall Rs value is 0.80 (p<.001), displaying a strong relationship between the two (Table 5). For dairies, there was an overall moderate correlation between area and total head of cattle, with a mean value of 0.50 (p<.001) with some regions with dairy farms displaying stronger correlations (Table 6).

Figures 7 and 8 display the relationship between feedlot or dairy area to total head of cattle. For all regions, as the total area increases, so does the total head of cattle. Feedlots have the largest R^2 values compared to dairies, ranging between 0.70 to 0.85 (p<.001). Additionally, feedlots tend to have larger cattle populations than dairies. In Figure 8, dairy R^2 values vary between 0.33 to 0.65 (p<.001). The slopes of the Midwest and China models tend to be less steep compared to the southwest dairy model. The lower R^2 values for dairies suggest there may be a density factor that occurs when housing dairy cattle. Based on our estimates pulled from academic reports describing facility design, the stall width and length requirements of dairy facilities tended to allocate less space to dairy cows, ~2.8m^2 to ~3.2m^2, compared to beef cows, ~2.3m^2 to ~14m^2 (Euken et al., 2015; McFarland and Tyson 2016; Krekelberg 2020). With less space allocated to dairy cows allows for more cows to be packed into a space. Based on the Spearman correlation analysis, and the scatter plots and R^2 values, these support our hypothesis that feedlot or dairy area can be used as a predictor of the total head of cattle.

Table 5 Spearman correlation coefficient between feedlot and dairy area to total head of cattle for different regions. N/A indicates there were no feedlots or dairies for comparison. The "***" represent statistically significant at the .01 level (*p*<.001).

Region	Feedlot area (ha)	Dairy area (ha)
Australia	0.70***	N/A
U.S., Michigan	N/A	0.41***
U.S., California	0.57***	0.68***
U.S., Texas	0.92***	0.90***
China	N/A	0.74***
South Africa	0.77	N/A
Overall	*0.80****	*0.50****

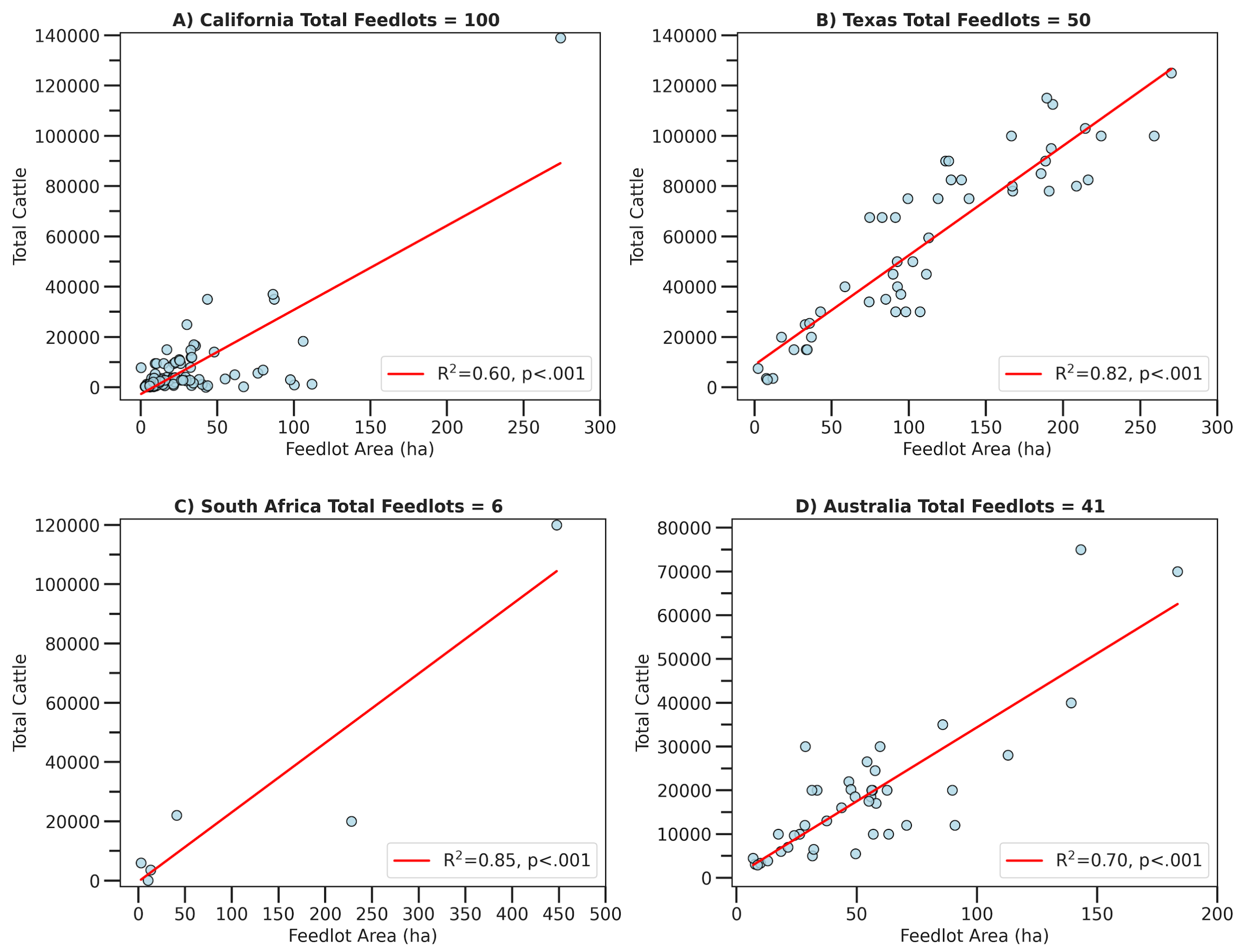

Figure 7 The relationship between feedlot area and total head of cattle for A) California, U.S., B) Texas, U.S., C) South Africa, and D) Australia. Number of feedlots, R^2 values, and statistical significance are included for each region. Note, the x- and y-axis ranges differ for each plot.

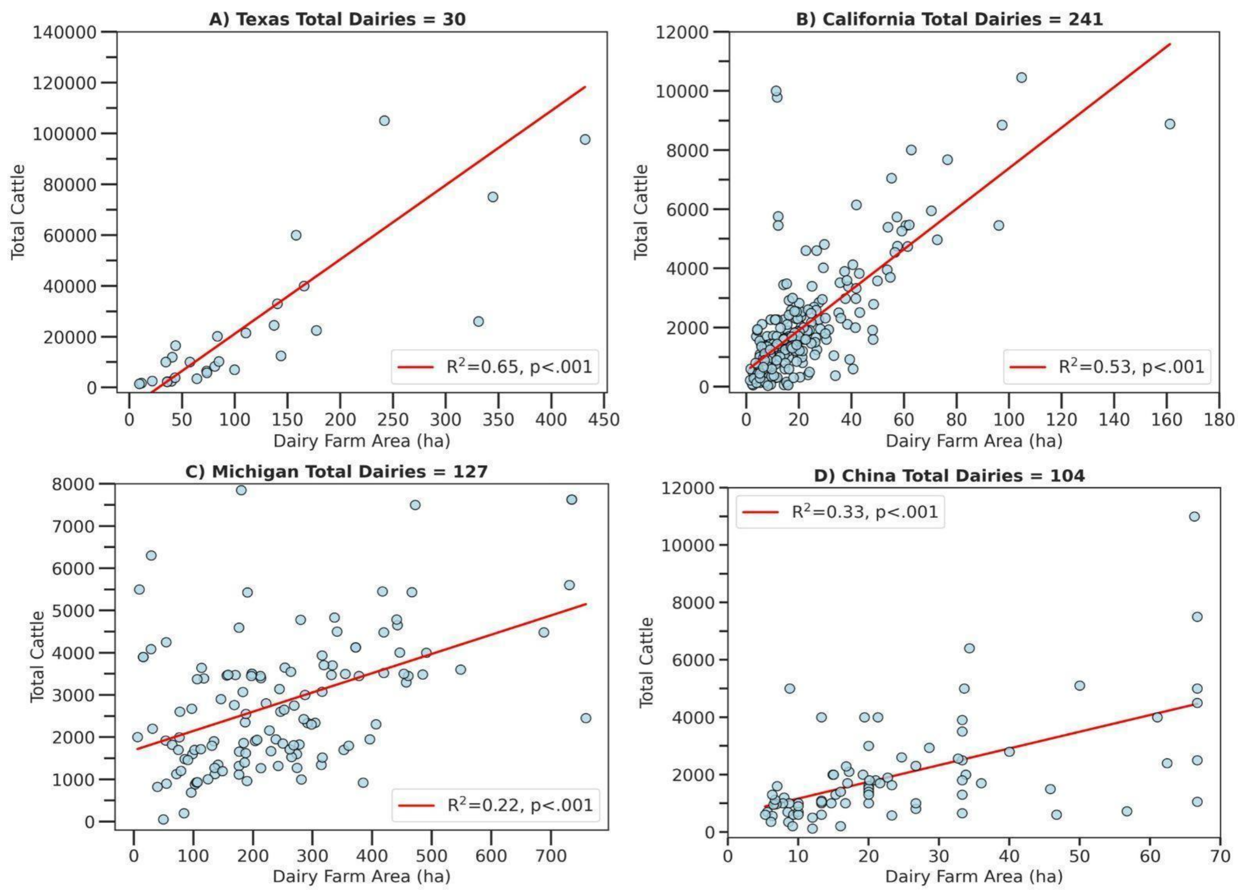

Figure 8 The relationship between dairy area and total head of cattle for A) Texas, U.S., B) California, U.S., C) Michigan, U.S., and D) China. Number of feedlots, R^2 values, and statistical significance are included for each region. Note, the x- and y-axis ranges differ for each plot.

3.1 Modeling results

Figure 9 reports each model's performance relative to permit/reported data. The *Feedlot U.S. Southwest* model has the largest RMSE at 13,115, followed by *Dairy U.S. Southwest* at 8,768, then *Dairy U.S. Midwest* at 1,386, and, lastly, *Dairy China at 1,453*. R^2 values range between 0.22 to 0.85 (p<.001). The *Feedlot U.S. Southwest* and *Dairy U.S. Southwest* model tend to underpredict when the total head of cattle is 55,000. Similarly, the dairy models underpredict when the total head of cattle is greater than ~2,000. All models overpredict when there is ~12,000 total head of cattle or less, for the southwest models, and less than 2,000 total head of cattle in the dairy model plots. The overprediction may be due to Climate TRACE seeking to identify the largest emitters in each sector. For this sector, we focused on the largest emitting cattle operations, as discussed in section 2.3.1, biasing the model training.

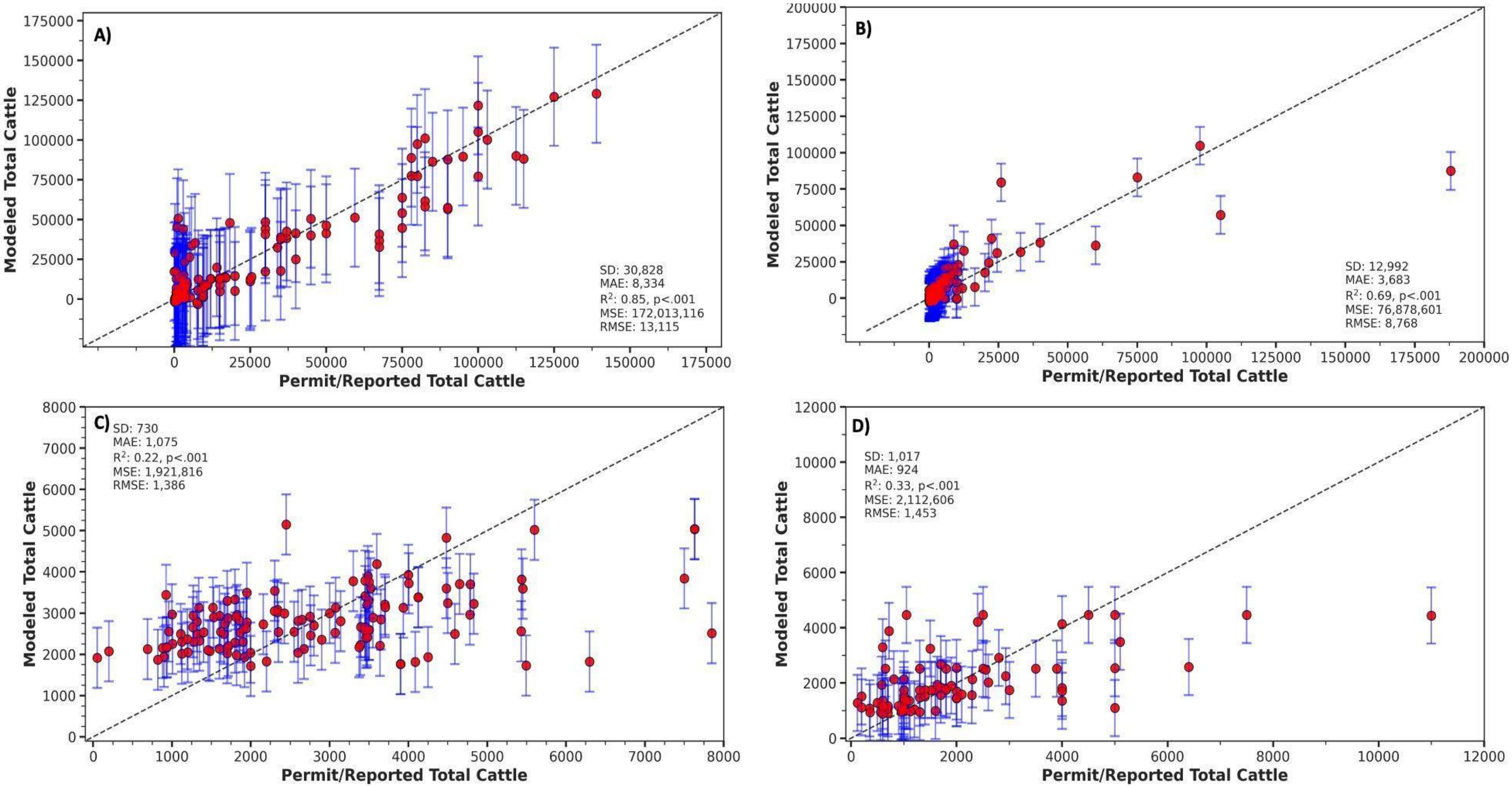

Figure 9 Scatter plots comparing permit/reported dairy population (x-axis) to estimated (modeled) cattle populations (y-axis). The models are as follows: A) *Feedlot U.S. Southwest*, B) *Dairy U.S. Southwest*, C) *Dairy U.S. Midwest*, and D) *Dairy China*. Blue vertical lines are 95% confidence intervals.

3.2 Emission estimates

Total CO_2e 100yr emissions estimates are shown for each region for years 2015 to 2022 in Figure 10. For the U.S., Texas has the highest emissions followed by California even though more cattle operations were identified in California (n=1,839) versus Texas (n=435). This can be due to Texas having some of the largest cattle operations identified in the Climate TRACE dataset, with the top four largest cattle operations are based in Texas. The largest Texas cattle operation contains 188,000 head of cattle according to Texas permit data (Table 1). The largest California cattle operation contains about half, ~97,000 head of cattle according to California permit data (Table 1). For the rest of the word, Argentina is the largest emitting country relative to the rest, minus the U.S, followed by Russia, then China. However, China has more identified cattle operations in our database (n=455) followed by Argentina (n=378), then Russia (n=186). The higher Argentinian emissions relative to Russia and China are due to the capacity utilization rates employed in each country, with Argentina having a higher value relative to China. Lastly, the U.S. has a larger emissions range related to other countries. This can be attributed to more U.S. cattle operations identified which have more detailed permit information that can provide specific manure management systems. This detailed information allows us to model emissions at an improved detail over other regions that lack such information.

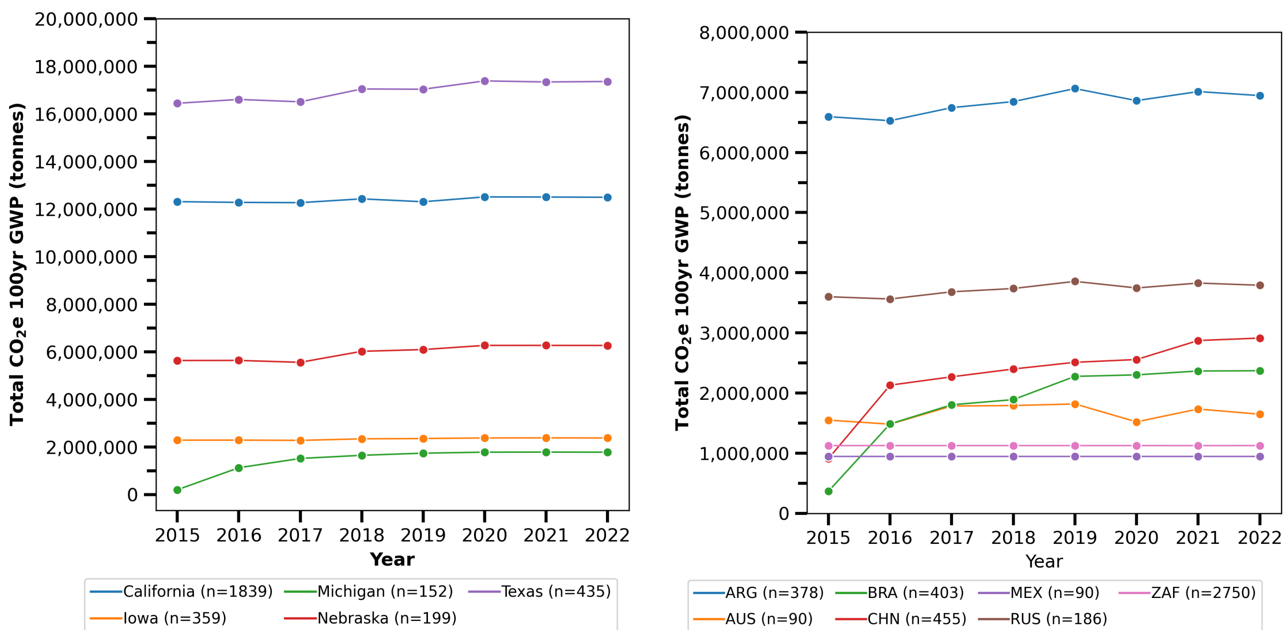

Figure 10 Time-series of total CO_2e 100-yr for U.S.: California, Texas, Michigan, Texas, Iowa, and Nebraska (left figure) and rest of the world: Argentina (ARG), Brazil (BRA), and MEX (Mexico), South Africa (ZAF), Australia, (AUS), China (CHN), and Russia (RUS) (right figure). Note the y-axis on each plot have different ranges.

A breakdown of the cattle population by type in each region is shown in Figure 11. Beef cattle are the dominant type in all regions. While the U.S. has permit information to identify the type of cattle held at each operation, such information is limited for the rest of the world. When identifying cattle operation types in other countries, we had to assume a default of "beef" when no discernable features were available to better identify the type of cattle operation.

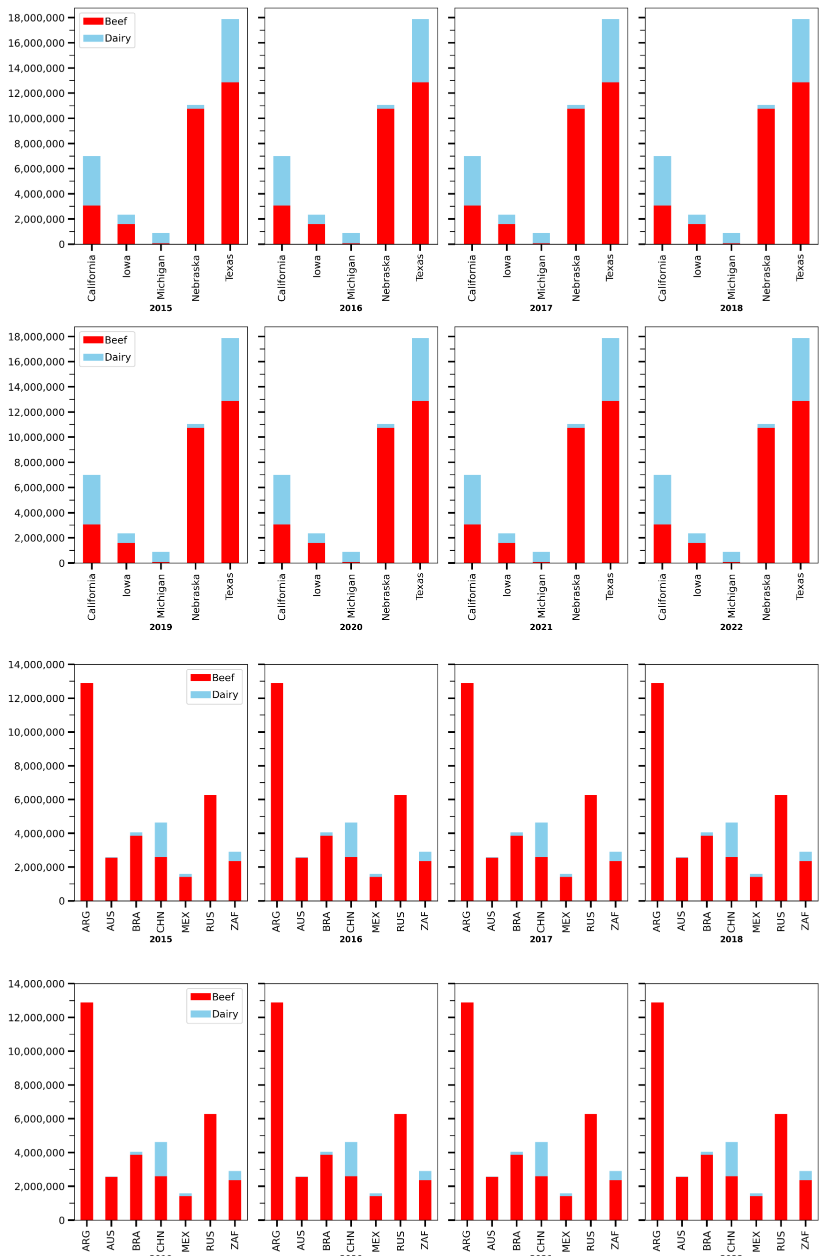

Figure 11 Stacked bar plots of total head of cattle broken down by beef (red bars) and dairy (blue bars) for U.S.: California, Texas, Michigan, Texas, Iowa, and Nebraska (top figure) and rest of the world: Argentina (ARG), Brazil (BRA), and MEX (Mexico), South Africa (ZAF), Australia, (AUS), China (CHN), and Russia (RUS) (bottom figure). Note the y-axis on each plot have different ranges.

Figures 12 to 14 provides examples of spatially map total aggregated feedlot and dairy emissions for years 2015 to 2022 for Texas, California, and China. Generally, based on the cattle operations Climate TRACE was able to identify ,each region generally has one dominant type of cattle: Texas has more feedlots than dairies, and California and China have more dairies than feedlots. Of interesting note is that dairy manure management emissions tend to have higher values relative to feedlot manure management emissions. This can be attributed to the manure management systems employed at dairies, which tend to be liquid/slurry based or aerobic treatment that produce higher emissions (Table S1). Generally, and assumed for feedlots, beef cattle operations use drylot manure management which produces less emissions relative to dairy systems.

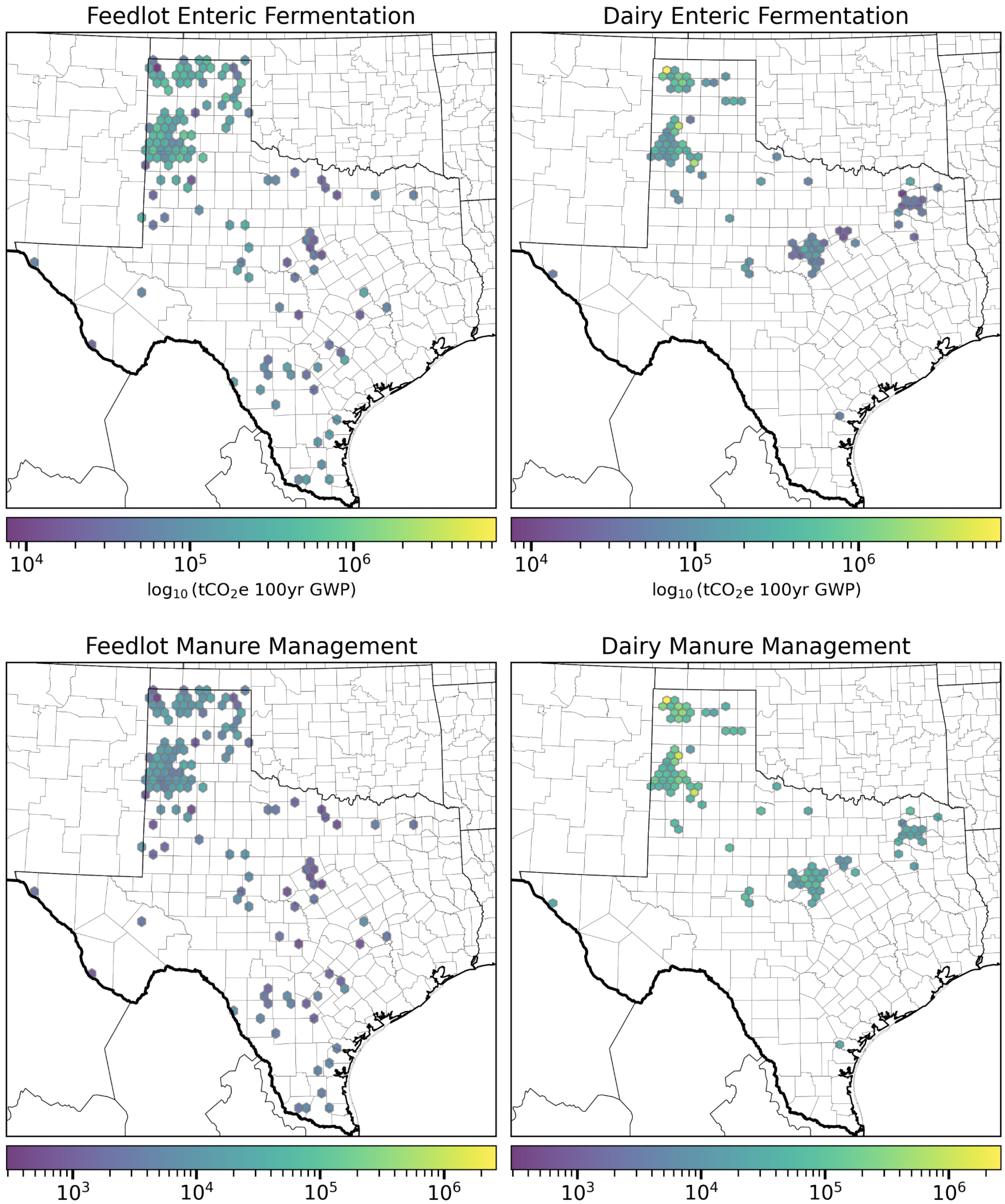

Figure 12 Texas spatially binned emissions for feedlots and dairies, represented as the total tCO_2e 100yr GWP emissions for years 2015 to 2022 (units in $\log_{10}$ scale). Top row displays enteric fermentation and bottom row displays manure management emissions for feedlots and dairies.

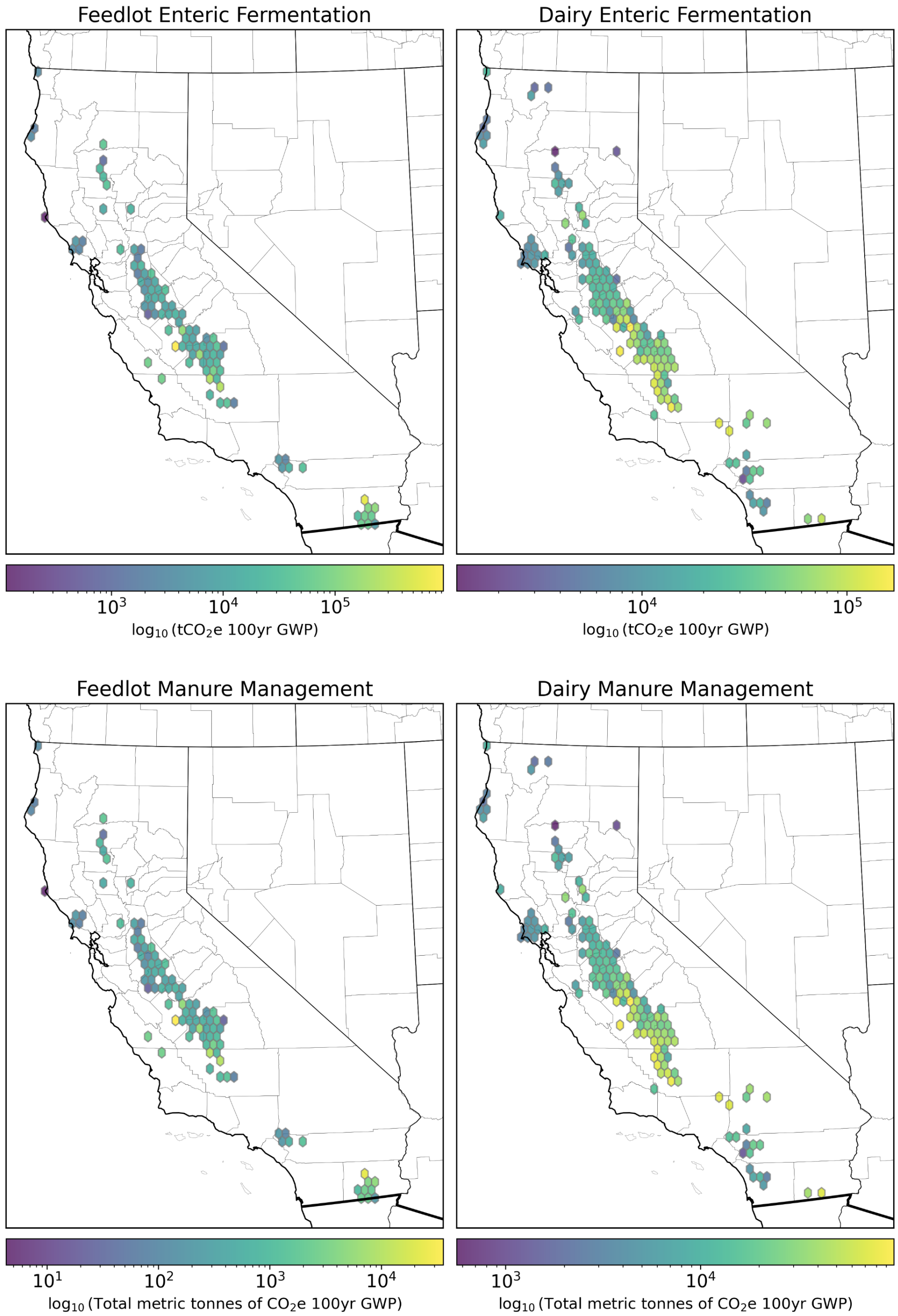

Figure 13 California spatially binned feedlots and dairies emissions, represented as the total tCO_2e 100yr GWP emissions for years 2015 to 2022 (units in $\log_{10}$ scale). Top row displays enteric fermentation and bottom row displays manure management emissions for feedlots and dairies.

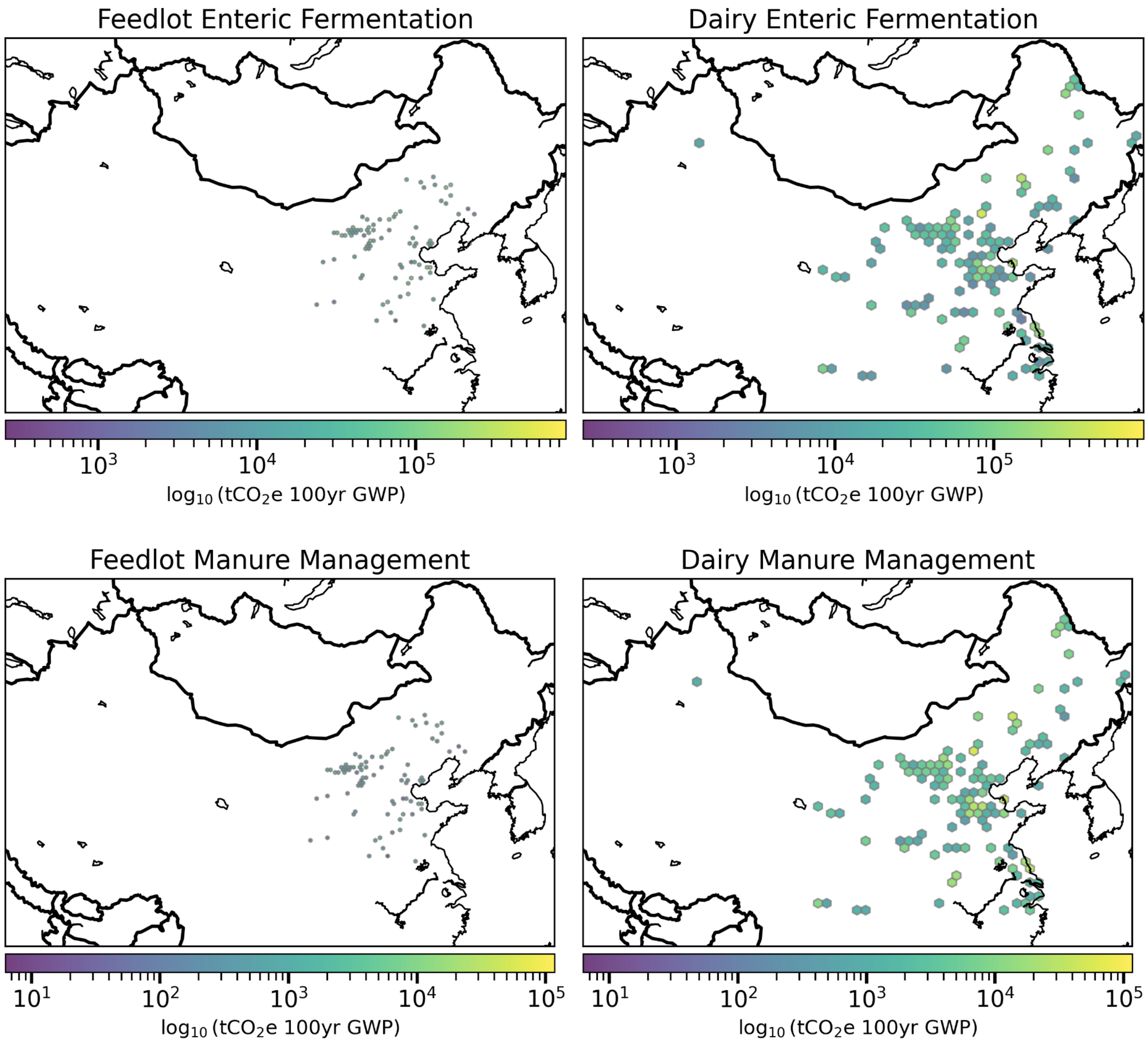

Figure 14 China spatially binned feedlots and dairies emissions, represented as the total tCO_2e 100yr GWP emissions for years 2015 to 2022 (units in $\log_{10}$ scale). Top row displays enteric fermentation and bottom row displays manure management emissions for feedlots and dairies.

4. Discussion and Conclusions

Climate TRACE's identification of feedlots and dairies and enteric fermentation and manure management emissions estimates provides insights of cattle operation emission sources. The deployment of the RAIC tool represents a scalable approach for spatially mapping and estimating individual feedlot and dairy emissions. This approach leverages available reported/permitting data with area estimates to estimate total head of cattle at individual cattle operations without publicly available population data. These paired approaches alongside the application of relevant emissions factors represents one of the first attempts to downscale emissions modeling to the facility level across an entire state or province.

There are conditions where this modeling approach works well for individual feedlots and dairies. RAIC was best at identifying cattle operations in contexts where they were sufficiently large and could be differentiated from nearby land uses. As a rule, RAIC is only capable of correctly labeling cattle operations in contexts where the human eye can successfully differentiate a feedlot from nearby land uses. Thus, small operations can blend into the landscape, or areas with low resolution imagery or high levels of haze that obscure aerial photography, extensive grazing lands that look similar to crop lands, or indoor facilities near adjacent enclosed buildings may be missed when utilizing the RAIC tool.

The Climate TRACE modeling approach works best when there are in situ, total head counts by cattle type at a representative portion of feedlots and dairies within a region, and where manure management can be identified. These conditions were not always present in feedlot and dairy producing regions, which may limit the applicability and precision of this approach. The population modeling requires some degree of local tuning to ensure that regional differences in management were captured, as shown in Figures 7 and 8. It was infeasible at this phase of the research to label the manure management practices of every cattle operation in the dataset or indicate where operations may have had mixed beef and dairy production systems. Operations were labeled either beef or dairy facilities, blurring the differences between integrated facilities. Differences in manure management between feedlots in a region were also difficult to capture in these results. All feedlots in a region were assigned one regional manure management practice (as were all dairies) if more detailed information was not available.

Future modeling will attempt to address several limitations of this process. Remote sensing approaches that detect features like manure lagoons, slurry pits, or anaerobic digesters, and label sites with the feature's accompanying manure management emissions factor may improve the model's performance in regions with more heterogeneous production systems. Feature identification will also improve the accuracy of feedlot and dairy area estimates. Figure 5 shows that some RAIC identified feedlots do not have their full area mapped. As a next step, RAIC will continue refining this approach. This includes providing satellite derived vegetation indices to separate crop areas from the convex hull drawing process and road feature identification to separate convex hull drawings that extend beyond a facility area.

The results reported on the Climate TRACE website results may be adjusted to account for differences in facility utilization, by comparing the estimated aggregate feedlot output (at 100% utilization) of identified livestock facilities to beef processing data and dairy production numbers. Adjustment factors may be incorporated should the estimates overestimate annual production. In areas where livestock spend less time on confined feedlots relative to open pasture, the current approach will fail to capture livestock emissions that originate upstream of feedlots in the supply chain. Countries with predominantly pasture-based systems, tend to still finish beef cattle on concentrated rations in feedlots even if cattle spend less time on those facilities, thus it may still be possible to reverse engineer emissions estimates from grazing operations by integrating the region's total feedlot capacity, the utilization rates of those facilities, the prevailing practices for the number of days cattle are on pasture before a feedlot, and the appropriate emissions factor for grazing in that area. Nonetheless, in places with smallholder cattle production systems that do not utilize feedlots, it will be very difficult to utilize remote sensing approaches like RAIC to assess cattle population and methane emissions. Mass balance or survey approaches may be the most appropriate available methods in those contexts.

The next step with this research will be to expand deployment of the RAIC model internationally. Climate TRACE only identified the largest cattle operations in each region. As a result, we only have a partially complete picture of the cattle operations present within a country. Climate TRACE will continually engage with Synthetaic to assess the viability of the RAIC tool in other dominant beef and dairy production countries, such as India and the European Union, and the remaining unmodeled regions of interest in the U.S., Russia, Argentina, China, Brazil, Mexico, and Australia. Climate TRACE will, in parallel, collect in situ, facility level population data to help tune the model to local practices, and test whether other variables may impact cattle density and warrant incorporation into the model. Future work may also integrate practice identification with remote sensing or even top down methane measurements, from providers such as AVIRIS and GHGSat, to help improve emissions modeling at identified facilities, or apply revised emission factors from Wolf et al. (2017) that account for breed and practice changes.

Acknowledgements
We would like to thank Jennifer Bockhahn (https://www.concordagriculturepartners.com/) for providing information on dairy and feedlot systems. Additionally, Matthew Scheffer, Lynn Henning, and Ann Marie Gardner for their engagement in this work through Hudson Carbon (hudsoncarbon.com) and Climate TRACE.

5. Supplementary data

Table S1 Permit/reported or manual manure management systems identified and IPCC equivalent used in emissions modeling. Included is the IPCC description of the manure management system, taken from IPCC (2006a).

Manure management system identified	**IPCC manure management system equivalent**	**IPCC description**
• Aerobic treatment • Lagoon - aerobic • Treatment & Storage pond • Treatment pond • Vegetative Infiltration Basin • Wetland	Aerobic treatment	The biological oxidation of manure collected as a liquid with either forced or natural aeration. Natural aeration is limited to aerobic and facultative ponds and wetland systems and is due primarily to photosynthesis. Hence, these systems typically become anoxic during periods without sunlight.
• Drylot	Drylot	A paved or unpaved open confinement area without any significant vegetative cover where accumulating manure may be removed periodically. Drylots are most typically found in dry climates but also are used in humid climates.
• Effluent Basin • Evaporation pond • Lagoon - anaerobic • Liquid slurry • Outside Concrete - uncovered • Outside Formed Concrete • Retention pond • Retention pond present • Runoff Control • Sand Settling Lanes • Settled Open Feedlot • Settling basin • Settling pond • Slurry Store • Solids Settling • Storage pond • Transfer pond	Liquid/slurry	Manure is stored as excreted or with some minimal addition of water in either tanks or earthen ponds outside the animal housing, usually for periods less than one year.

• Pit storage • Pit storage - Deep	Pit storage below animal confinements	Collection and storage of manure usually with little or no added water typically below a slatted floor in an enclosed animal confinement facility, usually for periods less than one year.
• Solid storage • Stockpiling Structure (covered or uncovered	Solid storage	The storage of manure, typically for a period of several months, in unconfined piles or stacks. Manure is able to be stacked due to the presence of a sufficient amount of bedding material or loss of moisture by evaporation.

5.1 Metadata section

Enteric fermentation and manure management emissions from individual cattle feedlots and dairies reports the following data on the Climate TRACE website:

- Individual feedlots and dairies enteric fermentation CH_4, and 20yr and 100yr GWPs emissions from feedlots and dairies
- Individual feedlots and dairies manure management CH_4 and N_2O emissions, and 20yr and 100yr from feedlots and dairies

Emissions estimates were reported for years 2015 to 2022. The cattle emissions described here represent a subset of specific country-level emissions estimates from the Climate TRACE agriculture sector: "*Country-level Enteric fermentation and Manure Management Emissions Estimates from Cattle Feedlots and Dairies*". Meaning", the country-level emissions encompass the subset of emissions contained in individual feedlot and dairy emissions estimates. This sector does not include cattle on pasture emissions. All data is freely available on the Climate TRACE website (https://climatetrace.org/). A detailed description of what is available is described in Table S2 to S4.

Table S2 Metadata for *Enteric Fermentation and Manure Management Emissions from Individual Cattle Feedlots and Dairies*.

General Description	**Definition**
Sector definition	*Individual feedlot and dairy emissions*
UNFCCC sector equivalent	*3.A.1 Cattle*
Temporal Coverage	*2015 – 2022*
Temporal Resolution	*Annual*
Data format	*CSV*
Coordinate Reference System	*None. ISO3 country code provided*

Number of emitters available for download	*12 countries total, representing 4,609 individual feedlots and dairies globally. For South Africa individual locations were aggregated to the province level, totaling 31 provinces with emission estimates*
Ownership	*Country*
What emission factors were used?	*IPCC CH. 10 and 11 EFs*
What is the difference between a "0" versus "NULL/none/nan" data field?	*"0" values are for true non-existent emissions. If we know that the sector has emissions for that specific gas, but the gas was not modeled, this is represented by "NULL/none/nan"*
total_CO2e_100yrGWP and total_CO2e_20yrGWP conversions	*Climate TRACE uses IPCC AR6 CO_2e GWPs. CO_2e conversion guidelines are here: https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_FullReport_small.pdf*

Table S3 Feedlot and dairy metadata description for confidence and uncertainty for enteric fermentation and manure management emissions.

Data attribute	Confidence Definition	Uncertainty Definition
type	• *Low*: if estimate is modeled or filled and location is similar to a feedlot or dairy • *Medium*: if estimate is modeled or filled and location found with research • *High*: if estimate type is permit/reported	Not used; N/A
capacity_description	• *Very low*: if estimate type was permit/reported or filled • *Medium*: if estimate type was modeled	Not used; N/A
capacity_factor_description	• *Very low*: feedlots, if capacity utilization rate is a fixed value or an average for all years (ex. ZAF, RUS, BRA) • *Low*: feedlots, if capacity utilization rate is a fixed value or an average for all years with a source (ex. CHN, MEX) • *Medium*: feedlot, if capacity utilization rate for different years was found (ex. USA, AUS); all dairies outside USA and AUS • *High*: All USA and AUS dairies	Not used; N/A
activity_description	• *Very low*: if estimate type was filled • *Medium*: if estimate type was modeled • *High*: if estimate type was permit/reported	• If estimate type was filled, standard deviation of feedlots less than 10 ha • If estimate type was modeled, standard deviation from the models

		• If estimate type was permit/reported, no uncertainty
CO2_emissions_factor	Not used; N/A	Not used; N/A
CH4_emissions_factor	*Medium*: based on IPCC emissions factors	IPCC uncertainty estimates, expressed as a percentage above or below the mean estimate (i.e. +/-XX%), or as an interval with an upper and lower bound of values.
N2O_emissions_factor	*Medium*: based on IPCC emissions factors	IPCC uncertainty estimates, expressed as a percentage above or below the mean estimate (i.e. +/-XX%), or as an interval with an upper and lower bound of values.
other_gas_emissions_factor	Not used; N/A	Not used; N/A
CO2_emissions	Not used; N/A	Not used; N/A
CH4_emissions	*Medium*: based on IPCC emissions factors	Given as an interval with an lower and upper bound of value
N2O_emissions	*Medium*: based on IPCC emissions factors	Given as an interval with an lower and upper bound of value
other_gas_emissions	Not used; N/A	Not used; N/A
total_CO2e_100yrGWP	*Medium*: based on IPCC emissions factors	Given as an interval with an lower and upper bound of value
total_CO2e_20yrGWP	*Medium*: based on IPCC emissions factors	Given as an interval with an lower and upper bound of value

Citation format: Davitt, A., Volpato, G., Gans, L., Cheng, X.F., Block, E., Raniga, K., Vandermel, J., Mendoza, A., McCrary, D., Sutherland, A., Rostami, R., Smith, M., Goodwin, B., Pluard, C., Schiller, S. (2023). *Enteric Fermentation and Manure Management Emissions from Individual Cattle Feedlots and Dairies*. WattTime, USA, Climate TRACE Emissions Inventory. https://climatetrace.org [Accessed date]

Geographic boundaries and names (iso3_country data attribute): The depiction and use of boundaries, geographic names and related data shown on maps and included in lists, tables, documents, and databases on Climate TRACE are generated from the Global Administrative Areas (GADM) project (Version 4.1 released on 16 July 2022) along with their corresponding ISO3 codes, and with the following adaptations:

- HKG (China, Hong Kong Special Administrative Region) and MAC (China, Macao Special Administrative Region) are reported at GADM level 0 (country/national);
- Kosovo has been assigned the ISO3 code 'XKX';
- XCA (Caspian Sea) has been removed from GADM level 0 and the area assigned to countries based on the extent of their territorial waters;
- XAD (Akrotiri and Dhekelia), XCL (Clipperton Island), XPI (Paracel Islands) and XSP (Spratly Islands) are not included in the Climate TRACE dataset;
- ZNC name changed to 'Turkish Republic of Northern Cyprus' at GADM level 0;
- The borders between India, Pakistan and China have been assigned to these countries based on GADM codes Z01 to Z09.

The above usage is not warranted to be error free and does not imply the expression of any opinion whatsoever on the part of Climate TRACE Coalition and its partners concerning the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its borders.

Disclaimer: The emissions provided for this sector are our current best estimates of emissions, and we are committed to continually increasing the accuracy of the models on all levels. Please review our terms of use and the sector-specific methodology documentation before using the data. If you identify an error or would like to participate in our data validation process, please contact us.

References

1. Costa Jr, C., Li, C., Cerri, C. E., & Cerri, C. C. (2014). Measuring and modeling nitrous oxide and methane emissions from beef cattle feedlot manure management: First assessments under Brazilian condition. *Journal of Environmental Science and Health, Part B*, *49*(9), 696-711.
2. Euken, R., Doran, B., Clark, C., Shouse, S., Ellis, S., Loy, D. and Schulz, L., 2015. Beef feedlot systems manual PM 1867. *Iowa State University, Ames, IA*.
3. Fan, X., Chang, J., Ren, Y., Wu, X., Du, Y., Xu, R., Liu, D., Chang, S.X., Meyerson, L.A., Peng, C. and Ge, Y., 2018. Recoupling industrial dairy feedlots and industrial farmlands mitigates the environmental impacts of milk production in China. *Environmental science & technology*, *52*(7), pp.3917-3925.
4. Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T.P., Vanwambeke, S.O., Wint, G.R. and Robinson, T.P., 2018. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific data*, *5*(1), pp.1-11.

5. Harper, L.A., Flesch, T.K., Powell, J.M., Coblentz, W.K., Jokela, W.E. and Martin, N.P., 2009. Ammonia emissions from dairy production in Wisconsin. *Journal of Dairy Science*, *92*(5), pp.2326-2337.
6. Harter, T., J. H. Viers, D. Liptzin, T. S. Rosenstock, V. B. Jensen, A. D. Hollander, A. McNally, A. M. King, G., Kourakos, E. M. Lopez, N. DeLaMora, A. Fryjoff-Hung, K. N. Dzurella, H. Canada, S. Laybourne, C., McKenney, J. Darby, and J. F. Quinn, 2013. Dairies and other sources of nitrate loading to groundwater. Task Report 6. SWRCB Agreement Number 04-184-555. Department of Land, Air, and Water Resources. University of California, Davis. 103 pages. http://groundwater.ucdavis.edu.
7. Hayek, M.N. and Garrett, R.D., 2018. Nationwide shift to grass-fed beef requires larger cattle population. *Environmental Research Letters*, *13*(8), p.084005.
8. Hopkins, F.M., T. Rafiq, and R.M. Duren. 2019. Sources of Methane Emissions (Vista-CA), State of California, U.S. ORNL DAAC, Oak Ridge, Tennessee, U.S. https://doi.org/10.3334/ORNLDAAC/1726
9. Iowa State University. 2018. *Iowa Farm Outlook: Capacity Utilization in the Cattle Feeding Sector Improving*. http://www2.econ.iastate.edu/ifo/files/ifo2018/ifo050118.pdf Accessed August 2022.
10. IPCC (Intergovernmental Panel on Climate Change), 2006a. Chapter 10. Emissions from livestock and manure management. *Guidelines for National Greenhouse Inventories. Vol. 4. Agriculture, Forestry and Other Land Use*.
11. IPCC (Intergovernmental Panel on Climate Change), 2006b. Chapter 11. N2O emissions from managed soils, and CO2 emissions from lime and urea application. *Guidelines for National Greenhouse Inventories. Vol. 4. Agriculture, Forestry and Other Land Use*.
12. *Major cuts of greenhouse gas emissions from livestock within reach*. FAO. (2013, September 26). Retrieved September 19, 2022, from https://www.fao.org/news/story/en/item/197608/icode
13. Krekelberg, E. 2020. *Space requirements for dairy cows.* https://extension.umn.edu/dairy-milking-cows/space-requirements-dairy-cows Accessed 30-Aug-2022.
14. McFarland, D. and Tyson, J. 2016. *Designing and Building Dairy Cattle Freestalls.* https://extension.psu.edu/designing-and-building-dairy-cattle-freestalls Accessed 20-Aug-2022.
15. Minerva Foods. 2020. *Sustainability Report 2020*. https://www.minervafoods.com/wp-content/uploads/2021/04/MinervaFoods_RS2020_EN.pdf Accessed 07 July 2023
16. Muñoz Sabater, J., (2019): ERA5-Land monthly averaged data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (<date of access>), doi:10.24381/cds.68d2bb30
17. Oklahoma State University (OSU), 2018. *Changes in Mexican cattle industry will affect American beef trade as well.*

https://news.okstate.edu/articles/agriculture/2018/changes-in-mexican-cattle-industry-will-affect-american-beef-trade-as-well.html Accessed 1 November 2023.

18. Owen, J.J. and Silver, W.L., 2015. Greenhouse gas emissions from dairy manure management: a review of field‐based studies. *Global change biology*, *21*(2), pp.550-565.
19. Peel, D.S., Mathews Jr, K.H. and Johnson, R.J., 2011. *Trade, the expanding Mexican beef industry, and feedlot and stocker cattle production in Mexico.* US Department of Agriculture, Economic Research Service, Outlook Report No. LDP-M-206-01. August. http://www. ers. usda. gov/media/118317/ldpm20601. pdf (Accessed 4 September 2023).
20. Pereira, H. 2022. *Technical Opinion: Critical evaluation of the methodology adopted in the IATP (Institute for Agriculture and Trade Policy) study to estimate JBS' GHG emissions.* https://jbs.com.br/wp-content/uploads/2022/08/-jbs-technical-note-iatp-eng.pdf
21. Accessed 07 July 2023
22. Petersen, S.O., 2018. Greenhouse gas emissions from liquid dairy manure: Prediction and mitigation. *Journal of dairy science*, *101*(7), pp.6642-6654.
23. Pienaar, J., Louw, A., and Jordaan, D. 2019. A Study on the potential product development for the commercialization and value add to beef products.
24. Skidmore, M.E., Sims, K.M., Rausch, L.L. and Gibbs, H.K., 2022. Sustainable intensification in the Brazilian cattle industry: The role for reduced slaughter age. Environmental Research Letters, 17(6), p.064026.
25. Turcios, S. 2022. *A breakdown of the American beef sector: How do U.S. cattle ranchers put food on our table?* Available at: https://clear.ucdavis.edu/explainers/how-do-us-cattle-ranchers-put-food-our-table Accessed: 27 September 2023.
26. U.S. Geological Survey. 2018. USGS EROS Archive - Aerial Photography - National Agriculture Imagery Program (NAIP) at https://www.usgs.gov/centers/eros/science/usgs-eros-archive-aerial-photography-national-agriculture-imagery-program-naip Accessed August, 2022.
27. U.S. Department of Agriculture (USDA), 2023. *Cattle Beef- Sector at a Glance*. Available at: https://www.ers.usda.gov/topics/animal-products/cattle-beef/sector-at-a-glance/ Accessed: 27 September 2023.
28. Waldrip, H.M., Todd, R.W., Parker, D.B., Cole, N.A., Rotz, C.A. and Casey, K.D., 2016. Nitrous oxide emissions from open‐lot cattle feedyards: a review. *Journal of environmental quality*, *45*(6), pp.1797-1811.
29. Waldrip, H.M., Parker, D.B., Miller, S., Miller, D.N., Casey, K.D., Todd, R.W., Min, B.R., Spiehs, M.J. and Woodbury, B., 2020. Nitrous Oxide from Beef Cattle Manure: Effects of Temperature, Water Addition and Manure Properties on Denitrification and Nitrification. *Atmosphere*, *11*(10), p.1056.
30. Waldron, S., Wang, J.M., Zhang, H.J., Dong, X.X. and Wang, M.L., 2015. The Chinese Beef Industry in "Regional workshop on beef markets and trade in Southeast Asian and China". *Ben Tre, Vietnam*.

31. Wolf, J., Asrar, G.R. and West, T.O., 2017. Revised methane emissions factors and spatially distributed annual carbon fluxes for global livestock. *Carbon balance and management*, *12*(1), pp.1-24.
32. Zahniser, S., Angadjivand, S., Hertz, T., Kuberka, L. and Santos, A., 2015. *NAFTA at 20: North America's free trade area and its impact on agriculture*. Washington, DC, USA: United States Department of Agriculture, Economic Research Service.
33. Zhu, G., Ma, X., Gao, Z., Ma, W., Li, J. and Cai, Z., 2014. Characterizing CH4 and N2O emissions from an intensive dairy operation in summer and fall in China. *Atmospheric Environment*, *83*, pp.245-253.