Emission Estimation Methods for Animal Feeding Operations

Draft

Prepared by:

U.S. Environmental Protection Agency

Office of Air and Radiation
Office of Air Quality Planning and Standards
109 T.W. Alexander Drive
Research Triangle Park, N.C. 27709

August 2021

Table of Contents

List of Tables

GLOSSARY / ACRONYMS

-2LogL	negative twice the likelihood
ADMs	average daily means
AFO	animal feeding operation
AIC	Akaike information criterion
AICc	adjusted Akaike information criterion
BIC	Schwarz Bayesian Information Criterion
FANS	Fan Assessment Numeration System
H_2S	hydrogen sulfide
LAW	live animal weight
MB	mean bias
ME	mean error
NAEMS	National Air Emissions Monitoring Study
NH_3	ammonia
NMB	normalized mean bias
NME	normalized mean error
PI	Principal Investigator
PM	particulate matter
PM_{10}	particulate matter with aerodynamic diameters less than 10 micrometers
$PM_{2.5}$	PM with aerodynamic diameters less than 2.5 micrometers
QAPP	quality assurance project plan
QC	quality control
TAN	total ammoniacal nitrogen
TEOM	tapered element oscillating microbalance
TKN	total Kjeldahl nitrogen
TSP	total suspended particulate
USDA	U.S. Department of Agriculture
VOCs	volatile organic compounds

1.0 INTRODUCTION

The EPA defines an animal feeding operation (AFO) as a lot or facility where: (1) livestock or poultry have been, are, or will be confined and fed for a total of 45 days or more in any 12-month period, and (2) crops, vegetative forage cover, or post-harvest residues are not sustained in the normal growing season over any portion of the lot or facility (40 Code of Federal Regulations, Section 122.23). The stipulation of the absence of vegetative cover intentionally excludes operations where animals are maintained on pasture or rangeland.

An AFO can include animal confinement structures, a system for managing manure, and might include a land application site for disposing of manure. Animal confinement structures may be totally enclosed with full-time mechanical ventilation (tunnel or cross ventilation), partially enclosed structures with or without mechanical ventilation, or paved or unpaved open lots. The manure management system collects manure, which includes a combination of fecal matter, urine, and other materials (e.g., bedding, waste feeds, wash water), and stores the manure until land application. Depending on the type of animal raised and the confinement method, manure can take the form of a solid, liquid, or slurry. Where the AFO handles manure as a solid, storage may occur within the confinement building (e.g., the manure accumulates on or beneath the flooring) or in covered or uncovered stockpiles. For liquid or slurry manure management systems, manure may be stored in an integral tank, such as a storage tank under the floor of a confinement building, or flushed to an external feature or structure, such as a pond or an anaerobic lagoon. Manure management systems can include stabilization processes to reduce volatile solids and odor emitted from manure prior to land application. Land application sites (e.g., cropland, pastures) typically serve as the ultimate disposal method for AFO manure. Solid manure can either be applied to the surface or applied to the soil surface followed by incorporation (e.g., disc harrowing). Liquid and slurry manure can be applied to the surface of soil, applied to the soil surface followed by incorporation, or injected into the soil. Table 1-1 presents an overview of the most common methods of confinement and manure management for large AFOs; however, the configuration and operation of individual AFOs varies depending on the animal type, regional climatic conditions, business practices, and operator preferences.

AFOs most common emissions include ammonia (NH_3), hydrogen sulfide (H_2S), volatile organic compounds (VOCs), and particulate matter (PM). These emissions have a variety of environmental and human effects. The compounds primarily responsible for the odors associated with AFOs are VOCs, H_2S, and other reduced sulfur compounds. VOCs also contribute to the formation of atmospheric ozone, which is a respiratory irritant. Some VOCs are designated in the Clean Air Act as hazardous air pollutants. Ammonia also is a source of odor from AFOs but to a lesser degree because NH_3 rapidly disperses in the air. Ammonia rapidly adheres to particles in the air due to its cohesive properties. Once released to the atmosphere, NH_3 is readily deposited

back to the earth in one of two forms, either gaseous ammonia (NH_3) or as the ammonium ion (NH_4^+). The ammonium ion can be converted to ammonium sulfate or ammonium nitrate, which can contribute to fine particulate concentrations ($PM_{2.5}$). When deposited back to the earth, these aerosols contribute to nutrient over-enrichment in aquatic systems and acidification of the environment.

Table 1-1. Common Types of Confinement and Manure Management Systems

Animal Type		Confinement Method	Manure Management System
Broiler chicken		Enclosed building	Integrated with confinement[a] or open or covered manure stockpiles
Egg-Layers	Dry manure		Integrated with confinement[a]
	Flush systems		Ponds and anaerobic lagoons
Swine			Integrated with confinement[a] or tanks, ponds, anaerobic lagoons
Dairy		Enclosed building and open lots	Anaerobic lagoons, tanks and ponds, and uncovered manure stockpiles

[a] The confinement building stores manure until land application.

2.0 AFO EMISSIONS

2.1 Gases

Gaseous emissions from AFOs largely result from the microbial decomposition of manure. Microbial decomposition of manure and the formation of gaseous compounds begin immediately upon excretion and continue until the manure is incorporated into the soil at the land application site. The composition of gas emissions and the subsequent emissions rates from AFOs depends on several factors: (1) the presence of an aerobic or anaerobic microbial environment, (2) the chemical precursors present in the manure, and (3) pH of the manure.

Gaseous emissions from AFOs are also affected by manure temperature and length of time before incorporation into soil. Temperature affects gas phase vapor pressure, and therefore, the volatility and emissions of water-soluble substances (e.g., NH_3), are greater at higher temperatures. Higher temperatures also favor the microbial processes that generate gaseous substances. Long periods of manure residence time in confinement, storage, or stabilization structures before incorporation into soil provide greater opportunities for anaerobic breakdown of manure and volatilization of gaseous emissions to the air. The method of applying manure can also affect emissions (e.g., emissions from land-applied manure not immediately incorporated will be higher than with rapid incorporation by disking or plowing).

2.1.1 Ammonia (NH_3)

Ammonia is a by-product of the microbial decomposition under both aerobic and anaerobic conditions of organic nitrogen compounds present in manure. The nitrogen compounds originate from unabsorbed nutrients in manure and in urine (urea from mammals and uric acid from poultry hydrolyze rapidly after excretion to form NH_3). Because NH_3 is highly soluble in water, NH_3 accumulates in manures handled as liquids and semi-solids or slurries but volatizes rapidly with drying and from manures handled as solids. Therefore, the potential for NH_3 volatilization exists wherever manure is present and NH_3 is emitted from confinement buildings, open lots, stockpiles, anaerobic lagoons, and land application from both wet and dry handling systems.

The volatilization of NH_3 from any AFO operation can be highly variable depending on total NH_3 concentration, temperature, pH, and storage time. Emissions will depend on how much of the ammonia-nitrogen in solution reacts to form NH_3 versus ionized ammonium (NH_4^+), which is nonvolatile. In solution, the partitioning of NH_3 between the ionized (NH_4^+) and un-ionized (NH_3) species is controlled by pH and temperature. Under acidic conditions (pH values of less than 7.0) ammonium is the dominant species, and NH_3 volatilization occurs at a lower rate than at higher pH values. However, some NH_3 volatilization occurs even under moderately

acidic conditions. Under acidic conditions, NH_3 that is volatized will be replenished due to the continual reestablishment of the equilibrium between the concentrations of the ionized and un-ionized species of NH_3 in solution following volatilization. As pH increases above 7.0, the concentration of NH_3 increases as does the rate of NH_3 volatilization. The pH of manures handled as solids can be in the range of 7.5 to 8.5, which results in fairly rapid NH_3 volatilization. Manure handled as liquids or semi-solids tend to have lower pH. Because of its high solubility in water, the loss of NH_3 to the atmosphere will be more rapid when drying of manure occurs. However, there may be little difference in total NH_3 emissions between solid and liquid manure handling systems if liquid manure is stored over extended periods of time prior to land application.

2.1.2 Hydrogen Sulfide (H_2S)

Microbial decomposition of sulfur contained in manure under anaerobic conditions produces H_2S and other reduced sulfur compounds (e.g., methyl mercaptan, dimethyl sulfide, dimethyl disulfide, and carbonyl sulfide). The sources of sulfur in animal manures are: (1) amino acids contained in the feed, (2) inorganic sulfur compounds (e.g., copper sulfate, zinc sulfate) used as feed additives to supply trace minerals and serve as growth stimulants, and (3) trace minerals in drinking water.

Manures managed as liquids or slurries are potential sources of H_2S emissions due to the anaerobic microbial environment. The magnitude of H_2S emissions is a function of liquid phase concentration, temperature, and pH. Temperature and pH affect the solubility of H_2S in water. The solubility of H_2S in water increases at pH values above 7. Therefore, the potential for H_2S emissions increases as the pH shifts from alkaline to acidic ($pH < 7$). Under anaerobic conditions, livestock and poultry manures will be acidic with pH values ranging from 5.5 to 6.5. At AFOs that handle manure in a dry form, the reduced sulfur compounds in manure are oxidized microbially to nonvolatile sulfate, and emissions of H_2S are minimal.

2.1.3 Volatile Organic Compounds (VOCs)

Volatile organic compounds are formed as intermediate metabolites in the degradation of organic matter in manure. Under aerobic conditions, any VOC formed are rapidly oxidized to carbon dioxide and water. Under anaerobic conditions, complex organic compounds are degraded microbially to volatile organic acids and other VOCs, which in turn are converted to methane and carbon dioxide by methanogenic bacteria. When the activity of the methanogenic bacteria is not inhibited, virtually all of the VOC are metabolized to simpler compounds, and the potential for VOC emissions is nominal. However, the inhibition of methane formation results in a buildup of VOC in the manure and ultimate volatilization to the air. Inhibition of methane formation typically is caused by low temperatures or excessive loading rates of volatile solids in

a liquid storage structure. Both of these conditions create an imbalance between populations of the microorganisms responsible for the formation of VOC and methanogenic bacteria. Therefore, VOC emissions will be minimal from properly designed and operated stabilization processes (such as anaerobic lagoons) and the associated manure application site. In contrast, VOC emissions will be higher from storage tanks, ponds, overloaded anaerobic lagoons, silage piles, and associated land application sites. The specific VOC emitted will vary depending on the solubility of individual compounds and other factors (including temperature) that affect solubility.

2.2 Particulates

The primary mechanism for direct releases of PM from AFOs is the entrainment of feeds, dry manure, soil, and other material (e.g., bedding materials, dander, feathers) caused by movement of animals in both indoor and outdoor confinement and activity at land application sites. The relative significance of each source depends on three interrelated factors: (1) the type of animal being raised, (2) the design of the confinement building being utilized, and (3) the method of manure handling (wet or dry). Additionally, precursors (e.g., NH_3 from AFOs and sulfur or nitrogen oxides) in the atmosphere can be converted to secondary particulate such as ammonium sulfate or ammonium nitrate.

Although all confinement buildings are sources of PM emissions, the composition of PM emissions varies with the only constant constituent being animal dander and feather particles from poultry. For poultry and swine, feed particles constitute a significant fraction of PM emissions because the dry, ground feed grains and other ingredients used to formulate these feeds are inherently dusty. Pelletizing of feeds reduces, but does not eliminate, dust and PM emissions. Dried forages also generate PM, but most likely to a lesser degree. Silages, which have relatively high moisture content, tend to generate less PM than for other types of feed. The mass of particulate matter emitted from totally or partially enclosed confinement buildings, as well as the particle size distribution, depend on type of ventilation and ventilation rate. Particulate matter emissions from naturally ventilated buildings will be lower than those from mechanically ventilated buildings. Mechanically ventilated buildings will emit more PM at higher ventilation rates. Therefore, confinement buildings located in warmer climates will tend to emit more PM because of the higher ventilation rates needed for cooling. While confinement buildings for dairy typically are all naturally ventilated, facilities for poultry and swine are mechanically ventilated for all or at least part of the year. When mechanical ventilation is used for only part of the year, it is used during the coldest and hottest months with natural ventilation used during the remainder of the year.

Open feedlots and storage structures for dry manure from broilers, turkeys, laying hens in high rise houses, and dairy drylots also are potential sources of particulate matter emissions. The rate of emission depends on whether or not the manure is covered. Open sites are intermittent sources of particulate matter emissions, because of the variable nature of wind direction and speed and precipitation. Thus, the moisture content of the manure and the resulting emissions will be highly variable. The PM emissions from covered manure storage structures depend on the degree of exposure to wind.

3.0 BACKGROUND OF EPA'S EFFORTS TO ESTIMATE AFO EMISSIONS

3.1 EPA's Air Compliance Agreement for Animal Feeding Operations

In August 2001, EPA published the report, *Emissions from Animal Feeding Operations,* which contains methodologies for estimating farm-level emissions from AFOs in the beef, dairy, swine, and poultry (broilers, layers, and turkeys) animal sectors (U.S. EPA, 2001). To develop the methodologies, EPA: (1) identified the manure management systems typically used by AFOs in each animal sector, (2) developed model farms, (3) conducted literature searches to identify emission factors related to model farm components (e.g., confinement, manure handling and treatment system), and (4) applied the emission factors to the model farms to estimate annual mass emissions.

After publication of EPA's 2001 report, *Emissions from Animal Feeding Operations,* EPA and the U. S. Department of Agriculture (USDA) jointly requested that the National Academy of Sciences (NAS) evaluate the current knowledge base and the approaches for estimating air emissions from AFOs. In a 2003 report, *Air Emissions From Animal Feeding Operations: Current Knowledge, Future Needs* (NRC, 2003), the NAS concluded the following: reliable emission factors for AFOs were not available at that time; additional data were needed to develop estimating methodologies; current methods for estimating emissions were not appropriate; and EPA should use a process-based approach to determine emissions from an AFO. Among the recommendations, was that in the short-term EPA should improve the then current emission factor approach while researching the implementation of a process-based model.

A process-based model is a mathematical representation of the biological processes occurring in a system. At its simplest, a process-based model traces the mass of an element through a biological process, ensuring the amount of that element leaving the system is consistent with the amount entering the system. With respect to AFOs, a process-based model would account for the nitrogen entering the system through feed, water, and the animals through the biological and chemical transformation that occur to the growing process; and ensuring this total mass equals the mass of nitrogen excreted in manure and urine, animal carcasses, and air emissions. As noted in the 2003 NAS report, process-based models are data intensive, requiring material sampling at all phases of animal development in addition to air monitoring. Therefore, EPA has proceeded with an approach to estimate emissions from sources based on statistical relationships between air emissions and the meteorological and housing parameters collected that are known to affect processes that generate emissions, as recommended for the near-term.

In January 2005, EPA announced the voluntary Air Compliance Agreement with the AFO industry. The goals of the Air Compliance Agreement were to reduce air pollution, monitor AFO

emissions, promote a national consensus on methodologies for estimating emissions from AFOs and ensure compliance with the requirements of the Clean Air Act (CAA), the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), and the Emergency Planning and Community Right-to-Know Act (EPCRA).

To develop the Air Compliance Agreement, EPA worked with industry representatives, state and local governments, environmental groups, and other stakeholders. Approximately 2,600 entities, representing nearly 14,000 farms that included broiler, dairy, egg layer, and swine operations, received EPA's approval to participate in the Air Compliance Agreement. As part of the Air Compliance Agreement, EPA agreed not to sue participating AFOs for certain past and ongoing violations of the CAA, CERCLA, and EPCRA, provided that participating AFOs agreed to pay a civil penalty (ranging from $200 to $100,000, based on the size and number of farms covered by a participant's Agreement) and comply with the Air Compliance Agreement's terms. The Air Compliance Agreement does not limit EPA's ability to act in the event of imminent and substantial danger to public health or the environment. The Air Compliance Agreement also preserves state and local authorities' ability to enforce local odor or nuisance laws. After EPA publishes the final emissions-estimating methodologies (EEMs) for the broiler, swine, egg layer and dairy sectors, participating AFOs must apply the final methodologies for their respective sectors to determine what actions, if any, they must take to comply with all applicable CAA requirements.

In the years since the Air Compliance Agreement, the Title XI of Division S of the March 2018 Consolidated Appropriations Act, also known as the Fair Agricultural Reporting Method Act (FARM Act), amended CERCLA to exempt air emissions from animal waste (including decomposing animal waste) at a farm from CERCLA reporting. Since that time, EPA has finalized rulemakings to provide a reporting exemption for air emissions from animal waste at farms from both CERCLA and EPCRA (84 FR 27533, June 13, 2019).

3.2 National Air Emissions Monitoring Study for AFOs

The AFOs participating in the Air Compliance Agreement were also responsible for contributing approximately $14.6 million to fund the National Air Emissions Monitoring Study (NAEMS). EPA provided oversight for NAEMS site selection and monitoring plans and for the team of researchers assembled from the following eight universities: Purdue University, Iowa State University, University of California-Davis, Cornell University, University of Minnesota, North Carolina State University, Texas A&M University and Washington State University. The NAEMS researchers conducted monitoring at 26 different confinement and open source sites considered to be representative of typical broiler, egg-layer, swine, and dairy operations in 10 states between 2007 and 2010, with data originally published by the NAEMS researchers in 2011

and finalized in 2012. Although not funded through the Air Compliance Agreement, the EPA considered a study conducted by Tyson Foods at two broiler farms in Kentucky (sites KY1B-1 and KY1B-2) from 2006 to 2007 to be an integral part of, and ultimately included in, the NAEMS dataset because the researchers at Iowa State University and the University of Kentucky (Burns et al, 2006) developed the quality assurance project plan (QAPP) for this study to be consistent with NAEMS QAPP. Table 3-1 lists the NAEMS monitoring sites by state.

Table 3-1. NAEMS Monitoring Sites

State	County	Site Name	Type of Animal and Operation Monitored
California	Stanislaus	CA1B	Broiler (Confinement)
California	San Joaquin	CA2B	Egg-Layer (Confinement)
California	San Joaquin	CA5B	Dairy (Confinement)
Iowa	Marshall	IA4B	Swine Sow (Confinement)
Iowa	Jefferson	IA3A	Swine Finisher (Manure Basin)
Indiana	Wabash	IN2B[a]	Egg-Layer (Confinement)
		IN2H[a]	Egg-Layer (Confinement)
Indiana	Carroll	IN3B	Swine Finisher (Confinement)
Indiana	Clinton	IN4A	Swine Sow (Manure Lagoon)
Indiana	Jasper	IN5B[b]	Dairy (Confinement)
Indiana	Jasper	IN5A[b]	Dairy (Manure Lagoon)
North Carolina	Nash	NC2B	Egg-Layer (Confinement)
North Carolina	Duplin	NC3B	Swine Finisher (Confinement)
North Carolina	Bladen	NC3A	Swine Finisher (Manure Lagoon)
North Carolina	Duplin	NC4A[c]	Swine Sow (Manure Lagoon)
		NC4B[c]	Swine Sow (Confinement)
New York	Onondaga	NY5B	Dairy (Confinement)
Oklahoma	Texas	OK3A	Swine Finisher (Manure Lagoon)
Oklahoma	Texas	OK4A[c]	Swine Sow (Manure Lagoon)
		OK4B[c]	Swine Sow (Confinement)
Texas	Deaf Smith	TX5A	Dairy (Confinement)[d]
Washington	Yakima	WA5A[c]	Dairy (Manure Lagoon)
		WA5B[c]	Dairy (Confinement)
Wisconsin	Saint Croix	WI5A[c]	Dairy (Manure Lagoon)[e]
		WI5B[c]	Dairy (Confinement)
Kentucky [f]	Union	KY1B-1	Broiler (Confinement)
	Hopkins	KY1B-2	Broiler (Confinement)

[a] Two different types of barns located at the same site were monitored.
[b] Monitoring occurred on two separate dairy farms in Jasper County, IN.
[c] Barns and lagoons were located at the same site.
[d] The reported emission estimates represent the entire corral.
[e] Instrumentation was deployed around two of the lagoons in the three-stage system. The emissions from the two lagoons were reported as a combined value.
[f] The Kentucky sites were part of an earlier Tyson Foods sponsored study that was designed to be consistent with NAEMS.

3.3 Previous Emission-Estimating Methodology Development

In February 2012, EPA developed draft EEMs for estimating air pollutant emissions from broiler confinement operations using the emissions and process data collected under NAEMS and other relevant information obtained through the 2011 *Call for Information*. The broiler draft EEMs included formulas to estimate NH_3, H_2S, and VOC emissions and emissions of coarse particulate matter (PM_{10}), fine PM ($PM_{2.5}$), and total suspended particulates (TSP). EPA also developed an EEM to estimate daily and annual NH_3 emissions from swine and dairy lagoons.

For broilers, EPA divided the process data into the following three groups: inventory (e.g., number of birds and bird weight), ambient (e.g., ambient temperature, pressure, and relative humidity), and confinement (e.g., building temperature, pressure, and relative humidity). The process parameters were statistically evaluated to determine if they were predictor variables. In addition, EPA evaluated whether the predictor variable process data were readily available to the growers, state and local agencies, and other interested parties. Given that the EEMs developed from the NAEMS dataset will be used for site-specific permitting decisions by operations participating in the Air Compliance Agreement, it is important to consider both the science in decision making, and the practical burden of collecting, maintaining, and supplying data to support emission estimations.

Based on the results of EPA's predictor variable evaluation process, three EEMs were developed using various process parameters. The three EEMs were: an EEM based on poultry inventory parameters (I EEMs); an EEM based on poultry inventory and ambient parameters (IA EEMs); and an EEM based on poultry inventory, ambient and confinement parameters (IAC EEMs). For the EEMs, EPA fit a polynomial mixed effects model (SAS version 9.2, Proc Mixed, SAS®) with an auto-regressive order 1 (AR(1)) covariance function.

At the time of its development, EPA explored the need for the model to include a random effect based on the house(s) monitored at each site, but EPA decided it was not needed. EPA also considered whether different variance values under different conditions were needed. The analysis suggested there was no evidence supporting an increase in the variance with increasing mean emissions. There was some indication of variance across the three sites, but it was decided not to include it in the model due to the limited data available.

EPA employed a backward elimination process to finalize the mean trend variables. To accomplish this, EPA started with an initial model run that included all variables (main effects and interactions). EPA then used the calculated p-values to eliminate the variables that were insignificant (p-value > 0.001). For this elimination process, the collection of cubic terms (e.g., [average mass, $(\text{average mass})^2$ and $(\text{average mass})^3$]) were considered the main

effect/interaction term. This collection of terms could only be removed as a group if a test of the null hypothesis that all three regression coefficients equal zero could be rejected. This was repeated until all terms remaining were significant (all p-values (or p-values for collections) < 0.001). At this point, EPA examined the fit statistics for the base dataset and cross-validation dataset to determine the version of the model (i.e., the set of mean trend terms) with the best performance. EPA selected this model as the candidate EEM.

EPA followed a similar process to develop an EEM to estimate daily and annual NH_3 emissions from the combined swine and dairy lagoon dataset. EPA used ambient temperature, relative humidity, solar radiation (represented by Julian day), and wind speed. Due to the very limited amount of data received for the nitrogen concentration, solid content, and pH of the lagoon liquid, those data were not included in the EEM.

In February 2012, EPA requested the SAB review and comment on these draft EEMs. Although the SAB reiterated that the models should be process based like the NAS 2003 report recommended, the SAB did acknowledge the NAEMS data were not sufficient to produce a process-based model. With respect to the statistical model itself, the SAB noted that EPA should not use a polynomial model because it leads to poor predictions at the extremes of the experimental conditions.

3.4 Current Emission-Estimating Methodology Development

EPA agrees that the development of process-based models should be pursued with the long-term goal of improving the accuracy of emission estimates for the livestock and poultry sectors. However, as noted in the SAB report, process-based models "require extensive data beyond the range of values, conditions, and types of farms available in NAEMS dataset." Given this current data limitation, EPA has developed statistical based models to serve as an initial emission estimation tool for AFOs to fulfill the requirements of the Air Compliance Agreement. These EEMs have built upon the recommendations of the SAB, to make these EEMs a better steppingstone toward the ultimate goal of process-based models.

Per the SAB recommendations, EPA adjusted the form of the modeling, variable selection, validation method used, and expanded residual analysis and evaluation of fit statistics. In this revised effort, EPA has dropped the use of polynomial forms to combat issues with extrapolation on the extreme ends of the data. In response to SAB comments and to move these statistical based models toward a process-based model, EPA has attempted to better represent the chemical, biological, and physical processes, and constraints in the EEM through the selection of variables used. EPA conducted a rigorous analysis of the literature and data available to identify the data elements collected under NAEMS with known chemical, biological, and physical

processes and constraints present at the monitoring sites. Those variables with the strongest connection to these processes were used in model development and selection was not completed strictly on significant p-values. For example, a primary driver of emissions is the volume of manure generated, as more manure has a higher emissions potential. The volume of manure generated is directly proportional to the number of animals present. Therefore, the inventory counts, or total live animal weight collected during NAEMS, are representative of this biological relationship.

With respect to the validation method, the previous efforts employed a k-fold cross-validation method, where 20 percent of the data were withheld to test the model. The SAB recommended splitting the data based on factors related to study design, such as house, to evaluate model predictive ability. In this current effort, EPA has shifted to a "jackknife" technique, which withholds each house at a time for model testing and validation.

For model fit evaluation, EPA expanded the use of residual diagnostic plots and also expanded fit statistics to include Akaike information criteria (AIC), adjusted Akaike information criteria (AICc) for number of predictors, Schwarz Bayesian criteria (BIC), and negative twice the likelihood (-2LogL) to measure predictiveness and effectiveness of fitted model. EPA also enhanced the model validation process, adding several standard statistics and metrics used throughout EPA to validate modeling.

4.0 AVAILABLE DATA

4.1 NAEMS 2010 Data (as published)

In the early planning stages of NAEMS, representatives from EPA, USDA, AFO industry, state and local air quality agencies, and environmental organizations met to discuss and define the parameters that would be collected by the study. The goal was to develop a comprehensive list of parameters that would provide a greater understanding and accurate characterization of the processes and activities at AFOs. By monitoring these parameters, EPA would have the necessary information to develop EEMs for uncontrolled emissions of particulate matter, NH_3, H_2S, and VOCs from animal feeding operations.

The NAEMS Protocol (Attachment B to the Air Compliance Agreement) provided guidance on the emissions and process parameters to be monitored under NAEMS and the specific components that were to be included in the emissions monitoring plans. In addition, the NAEMS Protocol identified the technologies and measurement methodologies to be used to measure emissions and process parameter data at each of the broiler, dairy, egg layer, and swine monitoring sites. The NAEMS Protocol required that an on-farm instrument shelter for housing monitoring equipment be located at each site and that the following parameters be monitored for 24 months:

- NH_3 concentrations using a chemiluminescence or photoacoustic infrared gas analyzer.
- Carbon dioxide (CO_2) concentrations using a photoacoustic infrared gas analyzer, or equivalent.
- H_2S concentrations using a pulsed fluorescence gas analyzer.
- $PM_{2.5}$ concentrations using a gravimetric, federal reference method for $PM_{2.5}$ for at least one month per site.
- PM_{10} concentrations using a tapered element oscillating microbalance (TEOM).
- TSP concentrations using an isokinetic, multipoint gravimetric method.
- VOC concentrations using a sampling method that captures a significant fraction of the 20 analytes determined by an initial characterization study of confinement VOC emissions to be the greatest contributors to total VOC mass (Heber et al, 2008).
- Animal activity, manure handling, feeding, and lighting operation.
- Total nitrogen and total sulfur concentrations determined by collecting and analyzing feed, water, and manure samples.
- Environmental parameters (heating and cooling operation, floor and manure temperatures, inside and outside air temperatures and humidity, wind speed and direction, and solar radiation).

- Feed and water consumption, manure production and removal, animal mortalities, and production rates.

The NAEMS Protocol also required sites to estimate the ventilation air flow rate of mechanically ventilated confinement structures by continuously measuring fan operational status and building static pressure, applying field-tested fan performance curves, and by directly measuring the air flow from selected fans using anemometers.

In addition, the NAEMS Protocol identified the technologies and measurement methodologies to be used to measure emissions and process parameter data at dairy and swine open source monitoring sites. The NAEMS Protocol required the use of optical remote sensing (ORS) techniques upwind and downwind of the lagoon combined with three-dimensional wind velocity measurements. The NAEMS Protocol required the following measurements:

- NH_3 and the various hydrocarbons concentration using open-path Fourier transform infrared spectroscopy (FTIR).
- H_2S and NH_3 concentration using collocated open-path ultraviolet differential optical absorption spectroscopy (UV–DOAS).
- Environmental parameters (air and lagoon temperatures, humidity, wind speed and direction, atmospheric pressure, and solar radiation).

The NH_3 and H_2S emissions were to be calculated from the difference in upwind and downwind concentration measurements using two different methods: a Eulerian Gaussian approach [computed tomography (CT)], and a Lagrangian Stochastic approach [backward Lagrangian stochastic method (bLS)]. For the VOC emissions, samples of the lagoon liquid were to be collected and analyzed for VOC, and EPA model WATER9 used to estimate emissions based on measured VOC concentrations, pH, and other factors.

There were some variations in parameters collected, because not all were applicable to each animal type and/or site. Additionally, some of the principal investigators (PIs) may have opted to collect more than required by the NAEMS Protocol.

To further supplement the NAEMS dataset, EPA published a *Call for Information* in January 2011 (76 FR 3060) to obtain emissions and process parameter datasets for animal confinement and manure storage and treatment operations at AFOs and supporting documentation. The *Call for Information* yielded 13 responses with reference to peer-reviewed journal articles or reports outlining other studies on AFO emissions. Because most of the data were not readily available in formats compatible with the NAEMS dataset, EPA used the data received from the *Call for Information* to inform decisions on parameter use. Additionally,

values reported in literature are also helpful for comparison with the current EPA draft EEMs to evaluate the reasonableness of these results.

4.1.1 Confinement Sites

The NAEMS collected a host of data from the sites including gaseous pollutant samples, particulate matter samples, meteorological data, confinement parameters, and biomaterial samples. All procedures were outlined in the project QAPP (Heber et al., 2008) and are summarized in the following section.

4.1.1.1 Emissions Sampling

NH_3 and H_2S concentrations were continuously sampled from multiple gas sampling probes with a custom-designed gas sampling system (GSS). Three gas sampling probes were placed in each barn in front of the exhaust fans. The inlet air (ambient air entering the barn) was sampled as well.

Each exhaust location was sampled individually for 10 minutes. The ventilation inlet location was monitored at least twice daily, originally with a 10-minute sampling period. The inlet sampling period was increased to 20 minutes and then 30 minutes later in the study.

Real-time PM monitors (TEOM Model 1400a, Thermo Fisher Scientific, Waltham, MA) were located immediately upstream of an exhaust in each house to continuously measure exhaust PM. A beta attenuation PM monitor (Beta Gage Model FH62C-14, Thermo Fisher Scientific, Franklin, MA) continuously measured house inlet PM concentration. At any one time, the sampled PM size class was PM_{10}, $PM_{2.5}$, or TSP at both TEOMs and the Beta Gage. The $PM_{2.5}$ size class was measured for at least twice over the course of the study, for a period of 4 to 21 days. The sites with less frequent $PM_{2.5}$ observation periods observed for longer periods. The TSP inlet heads were placed on the TEOMs for up to 10 times during the study duration, for 4- to 20-day periods. The PM_{10} concentration was measured at all other times.

The NAEMS PIs collected grab samples of VOC at the primary exhaust fans using methodology based on EPA Methods TO-15 (and TO-16). Sampling was conducted multiple times over the course of the study at each site, with duplicate samples typically collected at each location. All canisters were cleaned and passed quality control (QC) before sample collection. Canister samples were then analyzed at Purdue University's Trace Contaminant Laboratory. Samples were analyzed on a thermodesorption-gas chromatograph-mass spectrometer (TDS-GC-MS), consisting of a gas chromatograph (Model 6890, Agilent Technologies, Palo Alto, CA) coupled with a Model 5795 mass spectrometer detector (Agilent Model 5795) and equipped with a thermal desorption system (Model TDS-G, Gerstel, Baltimore, MD) and a cooled injection

system (Gerstel CIS). The analytical results were analyzed by ChemStation, and all integrations were manually checked. This method used an external standard compound for instrument monitoring and quality assurance (QA) to avoid losses of low-molecular-weight analytes that would occur when purging solvent used with internal standard(s). Response curves were generated at both the beginning and the end of the VOC analysis period.

4.1.1.2 Environmental Parameters

Building environment conditions were monitored through the study. Relative humidity and temperature (RH/T) probes were located at the primary representative exhaust fans (PREFs) for each barn. Additional thermocouples were used to measure temperatures inside the barns.

In-situ airflow measurements, or ventilation rate, were conducted with a 122-cm field-portable fan tester (Fan Assessment Numeration System (FANS), University of Kentucky, Lexington, KY), which was described by Gates et al. (2004). The field data were used to develop equations that would calculate airflow as a function of differential pressure and fan rotational speed, and to assess the uncertainty in airflow predictions.

Weather data were collected using a solar radiation shielded capacitance-type relative humidity and temperature probe (RH/T), a pyranometer, and a cup anemometer, which were attached to the roof of a barn or the instrument shelter installed for the study.

4.1.1.3 Animal Husbandry

Producers maintained records of mortalities, animal inventory, weight, and production (e.g., animals, eggs, milk), and water and feed consumption at each monitoring location. Producers also documented noteworthy activities or procedures that could affect emissions from each monitoring site (e.g., generator tests, manure removals, changes in diet and animal health, changes in ventilation, maintenance, cleaning, building cleaning, power failures).

4.1.1.4 Biomaterials Sampling Methods and Schedule

The NAEMS researchers periodically obtained manure samples from various storage or stabilization processes, including confinement structures with integrated manure storage, storage basins, lagoons, and storage piles using site-specific sampling equipment and techniques based on manure characteristics (i.e., liquid or dry samples) and facility configurations. Researchers collected composite samples comprised of numerous subsamples that were mixed (e.g., using a bucket or pail) to achieve a homogeneous state before sampling and analyzed the samples for pH, total and volatile solids content, and total nitrogen and NH_3 contents.

4.1.2 Open Sources

The NAEMS Protocol provided guidance on the emissions and process parameters to be monitored under NAEMS and the specific components that were to be included in the emissions monitoring plans. In addition, the NAEMS Protocol identified the technologies and measurement methodologies to be used to measure emissions and process parameter data at dairy and swine open source monitoring sites. The NAEMS Protocol required the use of optical remote sensing (ORS) techniques upwind and downwind of the lagoon combined with three-dimensional wind velocity measurements. The NAEMS Protocol required the following measurements:

- NH_3 and the various hydrocarbons concentration using open-path Fourier transform infrared spectroscopy (FTIR).
- H_2S and NH_3 concentration using collocated open-path ultraviolet differential optical absorption spectroscopy (UV–DOAS).
- Environmental parameters (air and lagoon temperatures, humidity, wind speed and direction, atmospheric pressure, and solar radiation).

The NH_3 and H_2S emissions were to be calculated from the difference in upwind and downwind concentration measurements using two different methods: an Eulerian Gaussian approach [computed tomography (CT)], and a Lagrangian Stochastic approach [backward Lagrangian stochastic method (bLS)]. For the VOC emissions, samples of the lagoon liquid were to be collected and analyzed for VOC, and the EPA model WATER9 used to estimate emissions based on measured VOC concentrations, pH, and other factors.

There were some variations in process parameters collected, because not all were applicable to each animal type or site. Additionally, some of NAEMS researchers opted to collect more data than required by the NAEMS Protocol. Table 4- lists the process parameters monitored at NAEMS open source sites and the open source project QAPP (Grant, 2008) outlines the data collection procedures.

Table 4-1. Continuous Parameters Monitored at NAEMS Lagoon Sites

Parameter		Units
Lagoon conditions	Temperature (lagoon liquid)	°C
	pH	pH
	Reduction/oxidation potential	millivolts
Meteorological conditions	Ambient temperature	°C
	Ambient relative humidity	%
	Barometric pressure	kPa
	Surface wetness	millivolts
	Solar radiation	Watts/m^2
	Wind speed	ft/sec
	Wind direction	Degrees

4.1.2.1 Emissions

Atmospheric concentrations of NH_3 around the basin were measured using narrow-bandwidth open path tunable-diode laser absorption spectroscopy (TDLAS). Atmospheric measurements of H_2S concentrations were made using pulsed fluorescence technology from air collected from 50-m synthetic open path systems (S-OPS) and sampled from a GSS that drew the air through the S-OPS. Emissions of NH_3 were determined from the difference in upwind and downwind concentration measurements from the TDLAS open path systems using two emissions models: a Gaussian plume fit model (Radial Plume Mapping: *RPM*; Arcadis Inc, Denver, CO) and a backward Lagrangian Stochastic (bLS) model (*WindTrax*; Thunder Beach Scientific, http://www.thunderbeachscientific.com).

Emissions of H_2S were determined using the concentration measurements from the pulsed fluorescence analyzer from air sampled by the air inlets of the S-OPS using two emissions models: a Ratiometric model, which uses the ratio of these concentrations to NH_3 concentrations along the same path with the corresponding RPM NH_3 emissions measurement, and the bLS model.

4.1.2.2 Weather Conditions

Measurements of the atmospheric temperature, relative humidity, barometric pressure, solar radiation, and surface wetness were measured and recorded at an automated weather station established on the basin rim.

4.1.2.3 Farm Activity

Additional information concerning farm operations was routinely collected from the producers. Pertinent activities affecting the basin include transfer of waste from barns and basin pump-outs for irrigation.

4.1.2.4 Basin Conditions and Biomaterial Sampling

For the basin site, the appearance of the basin was recorded on almost every site visit. Samples of the basin manure were collected during approximately each measurement period at the basin and analyzed for pH, total and ammoniacal nitrogen, sulfur, and total solids by a commercial laboratory. For the lagoon sites, measurements of the lagoon pH, oxidation-reduction potential (ORP), and temperature at 0.3 m depth were also measured from a float located at least 30 m from the lagoon inlet.

4.2 Revisions to 2010 Data Set

NAEMS PIs submitted monitoring data to EPA in 2010. These data are henceforth referred to as the "2010 dataset." More information about the QA associated with this dataset can be found in Grant et al. (2009) and Heber et al. (2008). The 2010 dataset was re-visited by Dr. Albert Heber (Barn PI and overall NAEMS PI) and Dr. Richard Grant (Open Source PI), who revised the barn and open source parts of the dataset, respectively. NAEMS PIs used this revised dataset for EEM development. The following sections provide a summary of the revisions for the barn and open source parts of the data.

4.2.1 Confinement Data

In early 2015, Dr. Albert Heber provided a revised swine barn dataset to EPA, which included a revision of the NH_3 and H_2S emission values due to a change in the methodology to determine barn gas inlet concentrations. NAEMS PIs revised the methodology to determine barn NH_3 and H_2S inlet concentrations by modifying the amount of time the concentration was valid to match the same amount of time used for the exhaust (outlet) concentration, which gave additional time for inlet gas concentrations to equilibrate from higher exhaust (outlet) concentrations. In addition, NAEMS PIs used a 10-day running average of inlet concentrations (5 days before and 5 days after) to determine NH_3 and H_2S emissions, which replaced the interpolated value between two individual measurements approximately 12 hours apart. This revision helped reduce the number of negative emission calculations due to occasionally high inlet concentrations. NEAMS PIs proposed that the 10-day running average was short enough to reflect seasonal changes in ambient concentrations, but also long enough to smooth out localized bursts at the inlet location. Further revisions included invalidating the air flow rate, and thus gas and particulate matter emissions, for periods when the ventilation was shut off.

4.2.2 Open Source Data

In late 2012, Dr. Richard Grant provided a revised open source dataset to EPA that contained 30-minute averaged emissions associated with the vertical radial plume mapping (VRPM) and bLS emission methodologies. When draft EEMs were first developed from 2010-2012, it was done without using the bLS data because the methodology had not been validated. However, the bLS method was validated in the Grant et al. (2013a) study, where it was evaluated in comparison to the VRPM method. In the Grant et al. (2013a) study, it was concluded that "The mean difference between the emission estimation methods is less than ±25% accuracy typically reported in the literature and consequently the two methods can be assumed to equally represent the actual flux conditions for open waste lagoon sources."

The 2012 bLS open source dataset had several revisions compared to the 2010 dataset. For the 2012 bLS open source dataset, NAEMS PIs applied a correction to all emission values, so they were reported at a standardized temperature of 20°C. NAEMS PIs also applied additional data validation criteria to the 2012 bLS open source dataset: the standard deviation of the wind direction had to be less than 30°C, the touchdown fraction had to be greater than 0.1, and, for NH_3, the background concentration had to between -0.1 ppm and 0.1 ppm. These criteria for valid data were also applied in studies by Grant et al. (2013a), Grant et al. (2016), and Grant and Boehm (2018), which reported and analyzed NH_3 emissions from NAEMS swine open sources. Additional information on the rationale for these criteria is provided in Grant et al. (2013a).

The 2012 VRPM dataset also had some additional data validation criteria compared to the 2010 dataset. NAEMS PIs applied additional criteria that the mean wind direction must be less than 60° of the perpendicular of the measurement plane and the upwind source fraction must be greater than 0.9. These criteria for data validity were also applied in the Grant et al. (2013a) and Grant et al. (2016) studies. Additional information on the rationale for these criteria is provided in Grant et al. (2013a).

Grant et al. (2013a) considered the bLS method to be closer (on average 5% less than VRPM in comparison) to the true emission value for lagoon sources since the VRPM method underestimates background concentrations by assuming that they are zero. Therefore, following the approach of Grant et al. (2016), NAEMS PIs adjusted lagoon VRPM NH_3 emissions for the individual site bLS-VRPM equivalency percentages reported in Grant et al. (2013a) and then averaged with bLS emissions. A different approach was used by Grant and Boehm (2018) for adjusting NH_3 emissions from a swine basin in Iowa. At this site, Grant and Boehm (2018) determined that the impact of nearby barn exhaust fans on measured NH_3 concentrations was likely to be greater in the bLS method than the VRPM method and thus bLS emission estimates were adjusted to VRPM emission estimates using an equivalency ratio. Accordingly, NAEMS PIs adjusted VRPM and bLS emission values and used these values to determine overall emissions for lagoon and basin emissions, respectively. If bLS and VRPM emission values were independent for a 30-minute time period, the individual adjustment, or non-adjusted VRPM/bLS, value was used; however, if both were valid, the two values were averaged.

In the 2010 dataset, NAEMS PIs reported H_2S emissions using the bLS methodology and the ratiometric emissions methodology. The ratiometric methodology is essentially the ratio of measured H_2S to NH_3 concentrations multiplied by the NH_3 emission rate. Thus, this calculation assumes that the same factors have the same influence on NH_3 emissions and H_2S emissions. This assumption is unlikely to hold for a variety of reasons, particularly since mass transfer processes govern NH_3 releases and ebullition processes control H_2S releases (Ni et al., 2009). Accordingly, only the bLS dataset was used for EEM development for H_2S. Furthermore, Dr.

Richard Grant has published two papers reporting H_2S from animal open sources (Grant et al., 2013b; Grant and Boehm, 2015), both of which only use bLS data.

5.0 PROCESS DESCRIPTIONS

5.1 Swine Operations

Swine operations breed and grow pigs for meat. Typical swine operations combine various stages of swine development. The number of swine sites in the United States has been steadily declining since 1959; however, the number of pigs marketed has increased. This is in part due to improvements in animal health (i.e., decrease in mortality rates) and increased sow fertility. It is also characteristic of the domestic swine industry becoming increasingly dominated by large totally enclosed confinement operations capable of handling 5,000 animals or more at a time.

The production cycle for swine has three phases: farrowing, nursing, and finishing. Some farms specialize in a single phase of the growth cycle, while other farms may handle two or all three phases. The first phase begins with breeding and gestation over a 114-day period followed by farrowing (giving birth). After farrowing, the newly born piglets normally are nursed for a period of three to four weeks until they reach a weight of 10 to 15 pounds. Sows can be bred again within a week after a litter is weaned. Sows normally produce five to six litters before they are sold for slaughter. After weaning, pigs are relocated to a nursery where swine typically are fed a corn-soybean meal based diet that may include small grains such as wheat and barley and other ingredients until slaughtered. Nursery operations receive weaned pigs and grow them to a weight of 40 to 60 pounds. The third phase of swine production is the growing-finishing phase where the gilts (young females) and young castrated boars (males) not retained for breeding are fed until they reach a market weight, typically between 240 and 280 pounds. Growing-finishing usually takes between 15 and 18 weeks, and animals normally are slaughtered at about 26 weeks of age.

Swine operations can be of several types with some farms specializing in a single phase of the growth cycle, while other farms may handle two or all three phases. As of the 2007 USDA Census of Agriculture (USDA, 2009), the most common is the growing-finishing operation, followed by the farrow-to-finish operation that encompasses all three phase of swine production. Another common production mode is the combination of the farrowing and nursing phases, which provide feeder pigs for stand-alone grow-finish operations. Although not as common, some newer farms may operate only the farrowing phase or only the nursery phase. These operations may be linked by common ownership or separately owned, but all under contract with a single integrator. Thus, pigs may begin their life-cycle in a sow herd on one site, move to a nursery on another, and then move again to a finishing facility. Specialized operations can take advantage of skilled labor, expertise, advanced technology, streamlined management, and

disease control. The farms that participated in NAEMS either encompassed the farrowing-nursing phases (also referred to as breeding and gestation farms) or finishing farms.

Barns for farrowing and finishing have difference concerns and management practices. Farrowing operations need intense management to reduce piglet mortality. Nursery systems are typically designed to provide a clean, warm, dry, and draft-free environment in which animal stress is minimized to promote rapid growth and reduce injury and mortality. Nursery buildings are cleaned and disinfected thoroughly between groups of pigs to prevent transmission of disease from one herd to another. Finishing pigs require less intensive management and can tolerate greater variations in environmental conditions without incurring health problems.

Four principal types of waste management systems are used with total and partially enclosed confinement housing in the swine industry: deep pit, pull-plug pit, pit recharge, and flush systems. (Other practices are used, but these are the predominant practices that swine operations currently use.) The deep pit, pull-plug pit, and pit recharge systems are used with slatted floors, whereas flush systems can be used with either solid or slatted floors. For flush systems, either fresh water or supernatant from an anaerobic lagoon transports accumulated wastes to an anaerobic lagoon. The pit may be flushed daily or as often as every two hours; the frequency depends on design characteristics such as channel length and slope and volume of water used per flush. In pit recharge systems, relatively shallow pits are drained periodically by gravity to an anaerobic lagoon. The frequency of draining varies but four to seven days is standard. Following draining, the empty pit is partially refilled with water, which can be supernatant from the anaerobic lagoon. Pull-plug pits are similar to pit recharge in that pit contents are drained by gravity to a storage or stabilization system. Pits are drained frequently, often each week or every two weeks. However, water is not added back into the pit. The system relies on the natural moisture in the manure. Deep pits are similar to pull-plug pits in that they store the manure directly under a slatted flooring system, and no water is added into the pit. They differ in that deep pits are typically sized to collect and store six months of waste. The accumulated manure has a higher solids content than pull-plug systems and emptied by pumping. To reduce odor, NH_3, and H_2S concentrations in confinement buildings with deep pits, ventilation air may flow through the animal confinement area, down through the slatted floor, and over the accumulated manure before discharge from the building. Alternatively, deep pits may be ventilated separately.

Most large swine farms have from 6 to 12 months of manure storage capacity (Pfost et al., 2000). Storage is in either an anaerobic lagoon or a storage structure. Typical storage structures include deep pits, tanks, and earthen ponds. Anaerobic lagoons provide both manure stabilization and storage. The use of storage tanks and ponds generally is limited to operations with deep pits and pull-plug pits where manure is handled as a slurry. Pit recharge and flush

systems typically use anaerobic lagoons, because of the need for supernatant for use as recharge or flush water

Each of these manure handling and storage methods affect emission of NH_3, H_2S, and VOC differently. Emissions of NH_3, H_2S, and VOC may be higher in flush systems than from pit recharge and pull-plug pit systems due to turbulence during flushing. Even with ventilation, emissions of NH_3, H_2S, methane, and VOC from confinement facilities with deep pits will likely be higher than from facilities with other types of manure collection and storage systems due to the sheer volume of manure stored.

5.2 Layers

Layers (defined as sexually mature female chickens capable of producing eggs (Wilson et al., 2000) typically enter the layer AFO at an age of approximately 18 weeks and stay in a layer house for approximately 18 months (UEP, 2017). Layer operations in the United States range in size from small farms of less than 20,000 birds to farms that are well over 100,000 birds. The largest facilities typically represent less than one percent of the total number of operations but confined over 55 percent of the laying hens (EPA 2001). Geographically, the most operations occur in the Midwest; however large production facilities are evenly spaced throughout the country.

The emptying of a flock from a house and the placement of a new flock can take 1-2 months. Commonly, the management practice of induced molting is used in layer operations, which is a management practice used primarily to optimize egg production and quality (Bell, 2003; Anderson, 2015). Furthermore, induced molting allows flocks to lay for additional time periods, thus reducing cost (Bell, 2003). Molting involves feed withdrawal, change of feed diet, and sometimes light exposure (Anderson, 2015). Molting typically lasts around 4-6 weeks (Anderson, 2015) and consists of three phases: a pre-molt phase, the weight loss phase (molt period), and a return to production phase (Anderson, 2015; Li et al., 2013).

Most U.S. layer housing types and manure management schemes fall under one of two categories. Facilities either utilize high-rise houses in which manure is stored in the lower level and removed every 1 to 2 years or belt houses with quasi-continuous manure transfer to an external storage/ treatment structure. In both housing arrangements, laying hens are almost exclusively confined in cages, which allow automation of feed distribution and egg collection.

5.3 Broilers

Broiler production is the raising of chickens of either sex for meat. A broiler is a young chicken that is characterized as having tender meat, flexible breastbone cartilage and soft, pliable

smooth-textured skin. Broiler production is a highly vertically-integrated industry, wherein a common owner or parent company is involved in several phases of the supply chain. For example, a parent company, or integrator, typically operates or contracts every aspect of the broiler production process (e.g., hatcheries, production houses, slaughterhouses, meat packing plants, feed production facilities and food distributors).

For broiler production operations, the integrator typically provides the birds, feed, medicines, transportation and technical support, under contract, to growers who provide the labor and the production facilities to raise the birds from hatchlings to market weight. The contract grower receives a minimum guaranteed price for the birds moved for market. More than 90 percent of all chickens raised for human consumption in the United States are produced by growers working under contract with integrators (USEPA, 2001). Because of this vertical integration, management strategies at the facility level tend to be more uniform than in other sectors of AFOs.

The length of the grow-out period ranges from 28 to 63 days, depending on the size of the bird desired. The grow-out period includes a brooding phase that begins when day-old chicks are placed in a heated section of a broiler house known as the brood chamber. The brood chamber is initially maintained at an elevated temperature (e.g., 85 to 95 °F), which is gradually decreased during the first few weeks of the birds' growth. As the growing birds need floor space, the remainder of the house is opened, and the chicks are grown to market weight.

Broilers are produced to meet specific requirements of customers, which can be retail grocery stores, fast-food chains, or institutional buyers. For broilers, the typical grow-out period is 49 days, resulting in an average bird weight of 4.5 to 5.5 pounds. The grow-out period may be as short as about 28 days to produce a 2.25 to 2.5 pound bird, commonly referred to as a Cornish game hen. For producing roasters weighing 6 to 8 pounds, the grow-out period is up to 63 days.

Broiler houses are operated on an "all in-all out" basis and require time for cleaning (e.g., decaking) and repair between flocks. For broilers, five to six flocks per house per year is typical; however, the number of flocks raised per year is lower for roasters and higher for Cornish game hens. Female broilers grown to lay eggs for replacement stock are called broiler breeders and are usually raised on separate farms. These farms produce only eggs for broiler replacements. A typical laying cycle for hens is about 1 year, after which the hens are sold for slaughter.

5.4 Dairy

<to be added>

6.0 OVERVIEW OF EEM DEVELOPMENT APPROACH

EPA developed EEMs separately for different swine production types (i.e., grow-finish, gestation, and farrowing) due to the significant differences in pig characteristics and the associated production and management conditions, which can have a large influence on air emissions. In addition, EPA developed separate EEMs for different open source waste management systems (i.e., lagoon and basin), because lagoons and basins have different storage times, which influence biochemical processes and thus air emissions.

The EEM development approach consisted of seven steps:

1. Data processing.
2. Trends analysis: Compare/contrast sites and review data plots to identify patterns.
3. Variable selection: Identify process and emissions variables for consideration.
4. Develop/refine/select daily emissions models.
5. Evaluation of daily emissions models.
6. Uncertainty estimates for annual emissions.
7. Model application.

6.1 Data Processing

The first step, data processing, consisted of loading and cleaning the dataset for use in the analysis. For data processing, the EPA-ORD standard operating procedure EMAB-129.0 : *Procedures for Entering or Importing Electronic Sample Data into Study Database* was followed. EPA imported the data from MS Excel® spreadsheets and MS Access database files into SAS®, a statistical analysis software package. Data was imported using a number of steps associated with the loading and where necessary, transposing of the data. EPA made only minimal adjustments to the dataset to ensure proper uploading into SAS, including adjustment to column names to comply with SAS string length and character limitations, and replacement of "not a number" (NaN) flags and Excel data errors flags (e.g., #VALUE, #N/A) with empty (null) cells. EPA reviewed the data to ensure appropriate transformations into SAS format.

Additional variables were also created in the SAS dataset by combining existing variables (e.g., live animal weight, which is a combination of animal inventory and animal weight) and adjusting existing variables for unit change (e.g., normalizing open source emissions for surface area). Variables were also added to the dataset to facilitate analysis by site, barn, date, or day of test. For open sources, additional data processing was needed to create ADM from 30-minute averaged data. A description of the method and data completeness criteria for determining ADM emissions is provided in Section 2 of the individual animal sector reports. For environmental data, a similar approach was used, with each 30-minute average considered valid if five or more of the six five-minute averages within the period were valid. ADM were calculated for

environmental variables if 36 or more values were valid in a day (out of 48 total), representing a completeness criterion of 75%.

The datasets were also updated in accordance with revisions that were made to the dataset as a result of identifying invalid data during the EEM development process. Section 2 of the animal sector reports have additional detail on specific revisions made to the datasets.

6.2 Trends Analysis

The second step of the EEM development process was to compare and contrast sites and identify patterns. The comparison of sites helped identify process differences that might be contributing to differences in emissions. For example, a site might have higher emissions values due to a different manure management system. This data exploration also helped to identify questionable data points for further review. This phase also included analysis to identify the strength of relationships between the available parameters and emissions. Section 4 shows the results of this analysis.

6.3 Parameters Selection

The third step identified the parameters to consider in EEM development. This step started with a literature review to identify parameters with established relationships with emissions. This was coupled with the exploratory data analysis to assess the strength of these well-established relationships within NAEMS data. The final phase evaluated the quantity of data available, the potential ease of variable measurement for a producer, and the exploratory data analysis together to select the variables to use in model development.

6.4 Daily model development

After the parameters were selected, EPA developed the daily models using a linear mixed effects model (SAS version 9.4, Proc Mixed, SAS®) to estimate average daily emissions at animal feeding operations by determining the effect of predictor variables on pollutant emissions. An advantage of using mixed models over standard linear models is that they allow for correlated errors, meaning that the mixed models can account for correlation among successive measurements. In this study, EPA accounted for correlation among successive measurements from each barn using a repeated variance spatial power covariance structure [Proc mixed SAS option: repeated day /subject=house type=sp(pow) day]. This covariance structure can be used when time intervals are not evenly spaced, which was a common occurrence in the NAEMS dataset due to missing data.

For modeling, all emissions were natural log transformed. To help with the log transformation, a constant (C) was added to the emission values of some EEMs (i.e., the same

constant value was applied to all emission values within an individual EEM) before log transformation to make all emissions values positive and/or not close to zero. This constant was subtracted from predicted emissions after back-transformation (see Section 7 of the individual animal reports). To avoid having several orders of magnitude difference between predictor variables, which can cause model convergence, units for live animal weight and inventory were changed to Mg and thousands, respectively. EPA's objective was to develop multiple models to predict average emissions in kg per day or grams per day, based on different combinations of the predictor variables (e.g., inventory, live animal weight, ambient temperature, and exhaust temperature) and then to evaluate the models based on their performance and how easily a producer could obtain measurements of the predictor values. When setting the combinations of the predictor variables to include, EPA often performed a pairwise correlation analysis to screen for predictor variables that might have a strong relationship to one another or could be linearly predicted from another variable. For example, correlation analysis found that live animal weight and inventory were highly correlated, which was expected because live animal weight is a function of inventory and average weight. Including related predictor variables in a multivariate regression can cause the estimates of the coefficient to change erratically in response to small changes in the data (e.g., when outliers are removed, or during model validation testing). Having related parameters as predictor variables does not affect how well the model can predict observations but can cloud the importance of any individual predictor. For this reason, strongly correlated parameters, such as live animal weight and inventory, were generally not included in the same model (except in some testing instances). Parameters with moderate correlations were used simultaneously in models because their interaction could be indicative of management practices. For example, the interaction of ambient temperature (an ambient parameter) and exhaust temperature (a barn parameter) could be informative of barn management practices.

For the particulate matter species (i.e., TSP, PM_{10}, and $PM_{2.5}$), EEM development started with the PM_{10} analysis, because PM_{10} measurements were the majority of particulate matter measurements taken over the course of the study. The PM_{10} dataset covered a broader set of meteorological and barn conditions. EPA considered how to ensure the more limited TSP and $PM_{2.5}$ datasets were still consistent with a broad range of conditions. For TSP, the literature reviews indicated similar emission processes responsible for the emission of both TSP and PM_{10}. Therefore, the parameters that influence PM_{10} should be similar to those that influence TSP, and a model that performs well for PM_{10} should also perform well for TSP. Therefore, the results for the PM_{10} analysis were considered when selecting a TSP model, giving preference to a model similar to one for PM_{10}, if all other parameters were equal.

Similarly, literature indicated that the processes for primary $PM_{2.5}$ in the barns would be similar to those for PM_{10}. $PM_{2.5}$ could be complicated by consideration of secondary formation

via chemical reactions. However, literature indicates that the formation of secondary $PM_{2.5}$ within the barns is probably minimal. EPA decided to consider the results for the PM_{10} analysis when selecting a $PM_{2.5}$ model, with preference given to models that included the same parameters, if evaluation statistics did not suggest otherwise.

EPA regressed the included predictor variables against natural log transformed average daily emissions, developed separate regressions for all combinations of the parameters, and eliminated from further consideration combinations that included insignificant predictors (p-value < 0.05). EPA assessed the fit of each model by preparing the residual diagnostic plots and used the following statistics to evaluate the predictiveness and effectiveness of the fitted models: Negative Twice the Likelihood (-2LogL), Akaike Information Criterion (AIC), Adjusted Akaike Information Criterion (AICc) for number of predictors, and Schwarz Bayesian Information Criterion (BIC).

Like the p-values, EPA calculated the values for -2LogL, AIC, AICc, and BIC from the "likelihood function" of an EEM, which quantifies the probabilities that different sets of parameter values will reproduce the emissions in NAEMS data. "Fitting the EEM" refers to finding the parameter estimates that maximize the likelihood function or finding the values of the parameters that account for the most variability in NAEMS data. Minimizing the function that is equal to -2LogL is mathematically equivalent to maximizing the likelihood, and the required computations take less time to perform. When comparing the values of -2LogL for two different EEMs, the model with the lower -2LogL value provides better fit to the data.

The equations below define the formulas for AIC, AICc, and BIC, where *d* is the number of model parameters and *n* is the sample size:

$$AIC = -2LogL + 2d \qquad \textit{Equation 1}$$

$$AICc = -2LogL + 2d\left(\frac{n}{n-d-1}\right) \qquad \textit{Equation 2}$$

$$BIC = -2LogL + \ln(n)\, d \qquad \textit{Equation 3}$$

All three criteria are functions of -2LogL, with an added term that penalizes inefficient models and models that achieve overfitting by adding more parameters. The equation for AIC is the simplest approach; the added penalty is twice the number of parameters used in the model ($2d$). The AICc and BIC statistics refine this approach by considering the overall sample size with respect to the number of parameters used in the model. AICc is generally considered the better statistic for small sample sizes. For all of these criteria, lower values indicate better performance of the model being evaluated relative to other models. EPA focused on AICc to

compare, rank, and select models that best explained the variation in NAEMS data, while penalizing candidate models that included greater numbers of predictors (Christensen et al., 2014).

EPA evaluated agreement between log-transformed observed emissions and model-predicted emissions using the following equation for log-transformed normalized mean error (LNME):

$$LNME = \frac{\Sigma|Y_p - Y_{lo}|}{\Sigma Y_o} \times 100\% \qquad \text{Equation 4}$$

Where:

Y_{lo} is the log transformed observed (or measured) emissions.
Y_p is the model predicted (log transformed) emissions.
Y_o is the observed (or measured) emission.

EPA assessed the agreement between the observed and predicted (back transformed from log) emissions for mean error (ME) and normalized mean error (NME) (defined below), which researchers have previously used in the evaluation of atmospheric and emission models (Walker et al., 2014; Rumsey and Aneja, 2014).

$$ME = \frac{\Sigma|Y_{bp} - Y_o|}{n} \qquad \textit{Equation 5}$$

$$NME = \frac{\Sigma|Y_{bp} - Y_o|}{\Sigma Y_o} \times 100\% \qquad \textit{Equation 6}$$

where Y_o is the observed (or measured) emission, n is the number of measurements, and Y_{bp} is the back transformed model-predicted log emission, using the following equation:

$$Y_{bp} = e^{\widehat{(y_p)}} * \overline{E}_i - C \qquad \textit{Equation 7}$$

Where:

Y_{bp} is the back transformed predicted emissions.
y_p is the model predicted (log transformed) emissions.
$\overline{E}_i$ is the average residual between model-predicted and observed (or measured) emissions on the natural log scale.
C is a constant added to the data prior to the log transformation.

The variable $\overline{E}_i$ includes an adjustment for bias associated with the log back-transformation (Newman, 1993). All EEMs expressed emissions as log transformed values. For back-transformed model-predicted emissions, Equation 5 should be used after the EEM calculations. The values of $\overline{E}_i$ and C for each EEM developed are provided in Section 7, along with an example calculation.

EPA assessed the agreement between the observed and predicted (back transformed from log) emissions for mean bias (MB) and normalized mean bias (NMB) using the following equations:

$$MB = \frac{\sum(Y_{bp} - Y_o)}{n} \quad \text{Equation 8}$$

$$NMB = \frac{\sum(Y_{bp} - Y_o)}{\sum Y_o} \times 100\% \quad \text{Equation 9}$$

Tables summarizing the models and the fit statistics are provided in Section 5 of the animal sector report.

To further evaluate the model fit, EPA developed scatter plots of the observed emissions versus the EEM predicted emissions. The plots include the one-to-one (1:1) line. Points that fall on the 1:1 line were predicted perfectly by the EEM. Points above the line indicates over predictions by the EEM (positive bias) and those below were under predicted by the model (negative bias). Plots for all the models tested are included in Appendix F or each animal sector report.

EPA also examined the models for outliers or questionable results (e.g., relationships that were contrary to those found in literature). If any were found, EPA explored refinements to the data that would result in improvements in model performance (e.g., removal of outliers, addition of other variables). EPA then reran the models with this refined dataset and repeated the review process. Once data refinement was complete, a final candidate model was selected for further evaluation.

6.5 Model Coefficent Evaluation

To ensure reliable prediction of the emissions, the coefficients of the candidate model were evaluated with the jackknife method (Christensen et al., 2014; Leeden et al., 2008), which examined the cumulative effect on coefficient estimates of multiple "minus-one" runs. The jackknife approach called for removing one of the independent sample units from the dataset. For NAEMS, the individual barns at each site and the monitored lagoons are the mutually exclusive independent sample units. EPA then determined the associated parameter estimates for the selected model based on this dataset. This was repeated for each of the sample units. These results were then compared to the model coefficients based on the full dataset (full model). For each jackknife model, the ME, NME, MB, and NMB were calculated, based on Equations 4 through 9, to facilitate comparison.

EPA also prepared plots showing the variation in coefficients and standard errors for the selected model and compared to each of the jackknife models. EPA interpreted these plots similar to the Tukey confidence interval plots in that, if the result for the jackknife model overlapped the results for the full model (i.e., the area highlighted in gray on the figures), then the model coefficients are not inconsistent with one another. If the omission of one monitoring unit (e.g., a barn or lagoon) resulted in a coefficient that was outside ± 1 standard error of the full model, the sample unit was reviewed to determine if a specific characteristic of that unit (e.g., animal placement strategy, manure handling system) might have caused the inconsistency. If the difference could not be ascribed to an operational characteristic of the unit, the data were reviewed for outliers that could be trimmed, and other potential remediation measures considered.

The evaluation statistics and plots for each candidate model is presented in Section 6 of each animal sector report.

6.6 Uncertainty Estimates

EPA also developed an estimate of uncertainty for total annual emissions, characterized by the random error in the model prediction using an approach similar to Monte Carlo analysis. Under this approach, EPA developed the statistical properties of predicted annual emissions by replicating annual sums of daily emissions. EPA ran these simulations for several different intervals of a predictor variable that fell within the observed range. For example, grow-finish barn testing live animal weight ranged from 0 to 130 thousand kg per head. The simulations were then run for live animal weight intervals of 3 thousand kg per head (e.g., 0, 3, 6, 9).

Simulations were run 10,000 times for each day for each interval to create an average uncertainty associated with the annual emissions from a single barn. EPA added a random residual to each day of the simulation to replicate the variability that would be seen in a real-world application of the model. For each of the intervals run, EPA calculated standard statistics (i.e., minimum, median, mean, maximum, range) and used these statistics to calculate the uncertainty for a single source, at that interval value, via Equation 39:

$$\text{Single source uncertainty} = 0.5 \times \left(\frac{\text{Range}}{\text{Median annual emission}}\right) \times 100 \quad \text{..... Equation 10}$$

EPA then plotted this single barn uncertainty against its associated annual emissions. This plot was then fit with a curve to model annual percent uncertainty for a single source (i.e., barn, house, lagoon, basin). For all uncertainty models, the curve took the form of:

$$Uncertainty\ (\%) = \frac{k}{Annual\ Emissions} \qquad \text{Equation 11}$$

Where k is a constant, and annual emissions are the total annual emission from the daily model.

Multiplying this percentage by the annual emissions calculated for the source provides the resulting uncertainty in the native emission units (e.g., kg or g), demonstrated in Equation 40.

$$Resulting\ Uncertainty = \frac{Percent\ uncertainty \times Annual\ emissions}{100} \qquad \text{Equation 12}$$

To propagate the uncertainty across all sources at a farm, EPA combined the estimates of absolute uncertainty for each source according to:

$$Total\ farm\ uncertainty = \sqrt{(U_{B1})^2 + \cdots + (U_{Bi})^2 + (U_{L1})^2 + \cdots + \left(U_{Lj}\right)^2} \qquad \text{Equation 13}$$

Where:

Total farm uncertainty = total uncertainty for the total emissions from all farm sources.
UBi = the resulting uncertainty for barns, and i represents the total number of barns on the farm,
ULi = the resulting uncertainty for open sources, and j represents the total number of open sources on the farm.

EPA notes that the uncertainty framework described above reflects the random uncertainty (error) in the prediction of daily emissions calculated using the EEMs, which includes the random uncertainty in the measurements used to develop the equation. This framework does not, however, consider systematic error (e.g., bias) in either NAEMS measurements or the EEM.

Section 7 of each animal sector report presents a summary of the uncertainty modeling and resulting equations.

6.7 Application

The final step is the application of the model to develop an emission estimate for a farm by a user. Section 8 of each animal sector report presents example calculations, including combining multiple structures for a to sum emissions on a farm and calculating an uncertainty estimate.

7.0 REFERENCES

Anderson, K. 2015. Induced Molting of Commercial Layers. *North Carolina State Extension*. AG-801. Raleigh, North Carolina.

Bell D.D. 2003. Historical and Current Molting Practices in U.S. Table Egg Industry. *Poultry Science Association Inc*. 965-970

Burns, R., Hongwei, X., Gates, R., Hoff, S., Nizich, S., Elkins, J. *Southeastern Broiler Gaseous and Particulate Matter Emission Monitoring*. (2006). Iowa State University, University of Kentucky.

Christensen J.R., Nash, M.S., Chaloud D., and Pitchford A. 2016. Spatial Distribution of Small Water Body Types Across Indiana Ecoregions. Ecohydrology. 9,122-137.

Grant, R.H., & Boehm, M.T. (2010a). *National Air Emissions Monitoring Study: Data from the southwestern US pork production facility OK4A*. Final report. Purdue University

Grant, R.H., & Boehm, M.T. (2010b). *National Air Emissions Monitoring Study: Data from the midwestern US pork production facility IN4A*. Final report. Purdue University.

Grant, R.H., Boehm, M.T., & Lawrence, A.F. (2013a). Comparison of a backward-lagrangian stochastic and vertical radial plume mapping methods for estimating animal waste lagoon. *Agricultural and Forest Meteorology 180,* 236-248.

Grant, R.H., & Boehm, M.T. (2015). Manure ammonia and hydrogen sulfide emissions from a western dairy storage basin. *Journal of Environmental Quality 44*, 127-136.

Grant, R.H., Boehm, M.T., & Lawrence, A.F. (2016). Ammonia emission from anaerobic waste lagoons at pork production operations: Influence of climate. *Agricultural and Forest Meteorology 228-229,* 73-84.

Grant, R.H., & Boehm, M.T. (2018). Ammonia emissions from an in-ground finisher hog manure tank. *Atmospheric Environment 190*, 43-52.

Heber A.J., Ni, J-Q., Ramirez, J.C., Schrock, W., & Elkins, J. (2008). *Quality assurance project plan for the National Air Emissions Monitoring Study (barn component)*. Purdue University.

Leeden (Van Der), Meijer R.E. and Busing F.M. 2007. Resampling Multilevel Models. Ch. 11. *Handbook of Multilevel Analysis*. Springer, New York.

Li Q-F., Wang-Li L., Bogan B.W, Wang K., Chai L., Ni J-Q., and Heber A.J. 2013. The National Air Emissions Monitoring Study's Southeast Layer Site: Part IV. Effects of Farm Management. *Trans. ASABE*. 56(3): 1199-1209.

MWPS (Midwest Plan Service). (1983). *Swine housing and equipment handbook*. Midwest Plan Service. *https://www-mwps.sws.iastate.edu/sites/default/files/imported/free/mwps8s.pdf*

Ni, J-Q., Heber, A.J., Sutton, A.L., & Kelly, D.T. (2009). Mechanisms of gas releases from swine wastes. *Transactions of the ASABE 52*(6), 2013-2025.

NRC (National Research Council). (2003). *Air emissions from animal feeding operations: Current knowledge, future needs.* National Academy Press. *https://www.nap.edu/catalog/10586/air-emissions-from-animal-feeding-operations-current-knowledge-future-needs*

Pfost, D.L., Fulhage, C.D., & Rastorfer, D. (2000). *Anaerobic lagoons for storage/treatment of livestock manure*. *https://extensiondata.missouri.edu/pub/pdf/envqual/eq0387.pdf*

UEP. 2017. Animal Husbandry Guidelines for U.S. Egg-Laying Flocks. Complete Guide for Cage and Cage-Free Housing. United Egg Producers. Johns Creek, Georgia.

USDA (U.S. Department of Agriculture) (2009). *2007 Census of Agriculture.* USDA, National Agricultural Statistics Service. *https://www.nass.usda.gov/Publications/AgCensus/2007/*

U.S. EPA (Environmental Protection Agency). (2001). Emissions from animal feeding operations (draft). http://epa.gov/ttn/chief/ap42/ch09/draft/draftanimalfeed.pdf

Wilson H.R., R.E. Buresh, F.B. Mather, W.D. Weaver. 2000. Poultry Glossary. Institute of Food and Agricultural Science. University of Florida. Gainesville, Florida.