Development of Emissions Estimating Methodologies for Lagoons and Basins at Swine and Dairy Animal Feeding Operations

Draft

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Sectors Policies and Programs Division
Research Triangle Park, NC 27711

February 2012

TABLE OF CONTENTS

Section **Page**

TABLE OF CONTENTS (continued)

Section **Page**

LIST OF TABLES

Table **Page**

LIST OF TABLES (continued)

Table **Page**

LIST OF FIGURES

Figure **Page**

Executive Summary

In 2005, the EPA offered animal feeding operations (AFOs) an opportunity to participate in a voluntary consent agreement referred to as the Air Compliance Agreement (Agreement) (70 FR 4958). Under the Agreement, participating AFOs provided the funding for the National Air Emissions Monitoring Study (NAEMS) – a two-year, nationwide emissions monitoring study of animal confinement structures and manure storage and treatment units in the broiler, egg-layer, swine, and dairy industries. The purpose of this study was to gather emissions data that would be used by the EPA to develop emissions-estimating methodologies (EEMs). The EEMs will be used by the AFO industry to estimate daily and annual emissions for use in determining their regulatory responsibilities under 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).

The NAEMS began in the summer of 2007 and consisted of 25 monitoring sites located in 9 states. Because the monitoring plan and quality assurance procedures were developed to be consistent with the NAEMS, the EPA also considered a monitoring study of two Kentucky broiler operations conducted by Tyson Foods from 2006 to 2007 to be an integral part of the NAEMS.

In accordance with the Agreement, the EPA developed EEMs for animal housing structures and manure storage and treatment units using the emissions and process data collected under the NAEMS and other relevant information submitted to the EPA in response to its Call for Information (76 FR 3060).

This report presents the background information, data collected, data analyses performed, statistical approach taken and the EEMs developed by the EPA for dairy and swine basins and lagoons. In the NAEMS documentation, the terms “lagoon” and “basin” were used inconsistently to describe the impoundments at the various monitoring sites. Although the EPA acknowledges that there might be differences between a lagoon and a basin (e.g., the degree of microbial activity), the term “lagoon” is used throughout this report to refer to lagoons and basins.

The EPA developed the EEMs using emissions and process information collected from nine lagoon monitoring sites across the country. Monitoring was conducted at three dairy farms located in Indiana, Washington and Wisconsin. Monitoring was also conducted at six swine farms: three breeding/gestation operations and three grow/finish operations. The breeding/gestation operations were located in Iowa, North Carolina and Oklahoma. The grow/finish operations were located in Indiana, North Carolina and Oklahoma. At the dairy and swine sites in Indiana, monitoring was conducted continuously for one year. The remaining seven sites were monitored for up to 21 days each season for two years by a team of researchers that moved sequentially from farm to farm.

The EPA used the emissions and process information collected at the nine study sites and SAS® statistical software to develop the NH_3 EEMs. Because of the limited number of some data values and gaps in coverage of seasonal meteorological conditions across monitoring sites, the EPA evaluated the combined dairy and swine data when developing the lagoon EEMs. This approach was taken to allow the effects of meteorological conditions on lagoon NH_3 emissions at a given site to inform the EEM for a site where meteorological data was limited. For example, limited data were submitted to the EPA for dairy lagoons at high ambient temperatures. By evaluating the dairy and swine data simultaneously, the EEM for dairy lagoons could be informed by the variability in swine emissions data that are available for the missing dairy lagoon temperatures.

The EPA developed three types of EEMs that include as predictor variables a combination of ambient meteorological data (e.g., temperature, relative humidity) that were continuously monitored and categorical (i.e., static) data that characterize the farm and lagoon configuration (i.e., animal type, farm capacity, lagoon surface area). The three EEMs resulted from the paired combinations of categorical farm and lagoon variables (i.e., animal type/surface area, animal type/farm size, farm size/surface area). Each EEM produces a point prediction (i.e., mean) of NH_3 emissions for a 30-minute period for a given set of input data. The animal type/farm size and animal type/surface area EEMs produce different emissions for swine and dairy lagoons. The EEMs also produce a 95-percent prediction interval to quantify uncertainty around the point prediction. The EEMs can be used to provide daily and annual estimates of NH_3 emissions from dairy and swine lagoons.

The EPA is currently developing EEMs for H_2S emissions. However, due to the very limited amount of H_2S emissions and process data available, the EPA is interested on obtaining feedback on the approach used for the NH_3 EEMs before completing development of the H_2S EEMs. Because VOC emissions data were not submitted to the EPA, EEMs for this pollutant were not developed. The EPA was unable to develop EEMs for VOCs, as the data for lagoon sources was not received.

1.0 INTRODUCTION

There are approximately 1 million livestock and poultry farms in the United States. About one-third of these farms raise animals in confinement, which qualifies them as Animal Feeding Operations (AFOs) (USDA, 2007). AFOs are potential sources of the following emissions: ammonia (NH_3), hydrogen sulfide (H_2S), total suspended particulate matter (TSP), particulate matter with aerodynamic diameters less than 10 micrometers (PM_{10}), PM with aerodynamic diameters less than 2.5 micrometers ($PM_{2.5}$) and volatile organic compounds (VOCs).

This report presents emissions-estimating methodologies (EEMs) for lagoons at swine and dairy sites. The methodologies were developed based on data collected in a National Air Emissions Monitoring Study (NAEMS) and other relevant information obtained through the EPA's January 19, 2011, Call for Information (see Section 3.0).

The EPA's previous effort to quantify potential emissions from this source sector and the evolution of the Air Compliance Agreement, are described in Section 1.1. Section 1.2 outlines the requirement for the NAEMS established by the Air Compliance Agreement. Section 1.3 describes how the data collected during the NAEMS was used to develop the EEMs.

1.1 Consent Agreement for Animal Feeding Operations

In August 2001, the EPA published methodologies for estimating farm-level emissions from AFOs in the beef, dairy, swine and poultry (broilers, layers and turkeys) animal sectors (*Emissions from Animal Feeding Operations*, Draft, August 2001). To develop the methodologies, the 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 the EPA's 2001 report, the Agency and the United States Department of Agriculture (USDA) jointly requested that the National Academy of Science (NAS) evaluate the current knowledge base and the approaches for estimating air emissions from AFOs. In its 2003 report (*Air Emissions From Animal Feeding Operations: Current Knowledge, Future Needs*, National Research Council), 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 the EPA should use a process-based approach to determine emissions from an AFO.

In January 2005, the 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, the EPA worked with industry representatives, state and local governments, environmental groups and other stakeholders. Approximately 2,600 AFOs, representing nearly 14,000 facilities that included broiler, dairy, egg layer and swine operations, received the EPA's approval to participate in the Air Compliance Agreement. Participating AFOs paid a civil penalty, ranging from $200 to $100,000, based on the size and number of facilities in their operations. They also contributed approximately a total of $14.6 million to fund the NAEMS.

As part of the Air Compliance Agreement, the EPA agreed not to sue participating AFOs for certain past violations of the CAA, CERCLA and EPCRA, provided that the AFOs comply with the Air Compliance Agreement's conditions. However, the Air Compliance Agreement does not limit the EPA's ability to take action 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 the 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, CERCLA and EPCRA requirements. If a participating facility ***does not*** trigger CAA, CERCLA or EPCRA permitting or release notification requirements based on the data collected, the facility will have 60 days from the publication date of the final EEMs to submit a written certification to EPA confirming compliance with current applicable requirements under these regulations. If a participating facility ***does*** trigger CAA, CERCLA or EPCRA permitting or release notification requirements, the facility will have 120 days from the publication date of the final EEMs to apply for any required permits under the CAA, or submit any required release notifications under CERCLA or EPCRA. Finally, AFOs that did not participate in the Air Compliance Agreement can use the appropriate EEMs for their sectors to determine what, if any, measures they must take to comply with applicable CAA, CERCLA and EPCRA requirements.

1.2 National Air Emissions Monitoring Study for AFOs

1.2.1 Overview of Emissions and Process Parameters Monitored

In the early planning stages of the NAEMS, representatives from the 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 must be monitored to provide a greater understanding and accurate characterization of emissions from AFOs. By monitoring these parameters, the EPA would have the necessary information to develop EEMs for PM, NH_3, H_2S and VOCs from animal feeding operations.

The Air Compliance Agreement provided guidance on the emissions and process parameters to be monitored under the NAEMS and the specific components that were to be included in the emissions monitoring plans. In addition, the Air Compliance Agreement 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 Air Compliance Agreement required the use of optical remote sensing (ORS) techniques upwind and downwind of the lagoon combined with 3-dimensional (3D) wind velocity measurements. The Air Compliance Agreement 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, as not all were applicable to each animal type or site. Additionally, some of the NAEMS researchers opted to collect more data than required by the Air Compliance Agreement. Table 1-1 shows the process parameters monitored at the NAEMS open source sites. Section 4.0 discusses the data received submitted to the EPA, including the amount of data received, in more detail.

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

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

1.2.2 NAEMS Monitoring Sites

Under the NAEMS, monitoring was conducted at 20 farms in nine states (California, Indiana, Iowa, New York, North Carolina, Oklahoma, Texas, Washington and Wisconsin). The study was conducted by a team of researchers 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. Consistent with the NAEMS Monitoring Protocol, the monitoring sites selected for the NAEMS provided representative samples of typical broiler, egg-layer, swine and dairy operations. The EPA provided oversight on site selection and monitoring plans. For the broiler sector portion of the NAEMS, Tyson Foods sponsored an earlier monitoring study at two broiler farms in Kentucky (sites KY1B-1 and KY1B-2) from 2006 to 2007. The quality assurance project plan (QAPP) for the Tyson study was developed to be consistent with the NAEMS. Therefore, the EPA considered the data collected at the Tyson study sites to be an integral part of the NAEMS. Table 1-2 presents the NAEMS monitoring sites.

For the open sources, monitoring was conducted at three dairies (IN5A, WA5A and WI5A), three breeding and gestation swine farms (IN4A, NC4A and OK4A) and three swine growing and finishing farms (IA3A, NC3A and OK3A). All farms operated under the overarching QAPP for the open source portion of the NAEMS (Appendix A).

Table 1-2. Monitoring Sites Under the NAEMS

State	County	Site Name	Type of Operation Monitored
California	Stanislaus	CA1B	Broiler (2 Houses)
California	San Joaquin	CA2B	Egg-Layer (2 High-Rise Houses)
California	San Joaquin	CA5B	Dairy (2 Barns)
Iowa	Marshall	IA4B	Swine Sow (2 Barns, 1 Gestation Room)
Iowa	Jefferson	IA3A	Swine Finisher (1 Lagoon)
Indiana	Wabash	IN2B[a]	Egg-Layer (2 Manure-Belt Houses)
		IN2H[a]	Egg-Layer (2 High-Rise Houses)
Indiana	Carroll	IN3B	Swine Finisher (1 "Quad" Barn)
Indiana	Clinton	IN4A	Swine Sow (1 Lagoon)
Indiana	Jasper	IN5B[b]	Dairy (2 Barns, 1 Milking Center)
Indiana	Jasper	IN5A[b]	Dairy (1 Lagoon)
North Carolina	Nash	NC2B	Egg-Layer (2 High-Rise Houses)
North Carolina	Duplin	NC3B	Swine Finisher (3 Barns)
North Carolina	Bladen	NC3A	Swine Finisher (1 Lagoon)
North Carolina	Duplin	NC4A[c]	Swine Sow (1 Lagoon)
		NC4B[c]	Swine Sow (2 Barns, 1 Gestation Room)
New York	Onondaga	NY5B	Dairy (1 Barn, 1 Milking Center)
Oklahoma	Texas	OK3A	Swine Finisher (1 Lagoon)
Oklahoma	Texas	OK4A[c]	Swine Sow (1 Lagoon)
		OK4B[c]	Swine Sow (2 Barns, 1 Gestation Room)
Texas	Deaf Smith	TX5A	Dairy (Corral)[d]
Washington	Yakima	WA5A[c]	Dairy (1 Lagoon)
		WA5B[c]	Dairy (2 Barns)
Wisconsin	Saint Croix	WI5A[c]	Dairy (2 Lagoons)[e]
		WI5B[c]	Dairy (2 Barns)
Kentucky[f]	Union	KY1B-1	Broiler (1 House)
	Hopkins	KY1B-2	Broiler (1 House)

[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, which were designed to be consistent with the NAEMS.

1.3 Emission-Estimating Methodology Development

The EPA developed an EEM to estimate daily and annual NH_3 emissions from swine and dairy lagoons. Section 5 describes the statistical methodology used to analyze the data and develop the EEMs. The EPA is currently developing EEMs for H_2S emissions. Because VOC emissions data were not submitted to the EPA, EEMs for this pollutant were not developed. The EPA was unable to develop EEMs for VOCs, as the data for lagoon sources was not received. Sections 3 and 4 discuss the NAEMS data received by the EPA to date.

The EPA used the continuous emissions, meteorological and lagoon data and categorical data collected under the NAEMS and SAS® statistical software to develop the NH_3 EEMs for swine and dairy lagoons. In the NAEMS documentation, the terms "lagoon" and "basin" were used inconsistently to describe the impoundments at the various monitoring sites. Although the EPA acknowledges that there might be differences between a lagoon and a basin (e.g., the degree of microbial activity), the term "lagoon" is used throughout this report to refer to lagoons and basins.

All of the NAEMS meteorological, lagoon and categorical data were assessed for data completeness and statistically evaluated to determine if they were reasonable predictor variables. In addition, the EPA evaluated whether the predictor variables that would be used in the EEM would be readily available to the growers, state and local agencies and other interested parties.

The EEM input parameters based on the continuous data are the 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, these data were not included in the EEM. The limited number of sites monitored under the NAEMS also limited the number of categorical (i.e., static) variables that could be included in the EEM without compromising the performance of the methodology to two variables. The EPA developed three EEM variations to evaluate the paired combinations of the following three categorical variables that were expected to have the greatest influence on NH_3 emissions from lagoons: animal type, farm size and lagoon surface area.

2.0 OVERVIEW OF OPEN SOURCES

As previously mentioned, the goal of the NAEMS was to quantify emissions from animal feeding operations. This included quantifying emissions from confinement sources (e.g., barns) and open sources (i.e., lagoons and corrals). For this report, manure is defined as any combination of fecal matter, urine and other materials that are mixed with manure (e.g., bedding material, waste feeds, wash water). Manure can be in a solid, slurry, or liquid state (e.g., surface liquids from storage facilities). The state of the manure often dictates the management practices and the degree to which pollutants are emitted.

Section 2.1 provides an overview of the general practices of the dairy and swine industry to provide context to study decisions (e.g., the types of farms to monitor). Section 2.2 describes the manure handling practices of the dairy and swine industry. Section 2.3 provides typical emissions from open sources. Section 2.4 describes the dairy and swine sites monitored under the NAEMS.

2.1 Overview of Dairy Industry

Dairies are AFOs that produce milk, raise dairy replacement heifers or raise calves for veal. Typically, dairy operations combine milk production and the raising of heifers (immature females) as replacements for mature cows that no longer produce milk economically. However, some milk producers obtain some or all replacement heifers from stand-alone heifer operations. Although some dairies raise veal calves, veal production is typically specialized at operations solely raising veal calves. For several decades, the number of milk producing cows has steadily decreased while the volume of milk produced has continually increased. This increased productivity has been the result of improvements in breeding programs and feeding and management practices.

Concurrently, there has been an ongoing consolidation in the dairy industry resulting in fewer but larger farms. Between 1991 and 2007, the number of dairy farms decreased by 58.4 percent, while milk production increase by 23.1 percent. As of the 2007 USDA census of Agriculture, the states with the largest number of dairy operations are Wisconsin (14,900), Pennsylvania (8,700), New York (6,400) and Minnesota (5,400).

2.1.1 Production Cycle

The primary function of a dairy is the production of milk, which requires a herd of mature dairy cows that are lactating. In order to produce milk, the cows must be bred and give birth. The gestation period is nine months, and dairy cows are bred again four months after calving. Thus, a mature dairy cow produces a calf every 12 to 14 months. Therefore, a dairy operation will have several types of animal groups present, including calves, heifers, mature cows (lactating and dry cows), veal calves, and bulls. The production cycle in the dairy industry begins with the birth of calves which causes the onset of lactation (milk production). A period of between 10 and 12 months of milk production is followed normally by a two-month dry period. The dry period allows for physiological preparation for the next calving. At the time milking normally is stopped, a cow normally will be in the seventh month of a nine month pregnancy. A high frequency of calf production is necessary to maintain a cost-effective level of milk production. The rate of milk production peaks shortly after calving and then slowly declines with time.

2.1.2 Animal Confinement

How dairy cows are confined depends on the size of the operation, age of the animal, and the operator preference. Dairies predominantly use some type of multiple animal area for unweaned calves, weaned calves, and heifers. Mature cows, when not being milked, are confined in freestall barns, drylots, tie stalls/stanchions, pastures, or combinations of these. Dry cows are confined in loose housing or freestalls.

Lactating cows require milking at least twice per day and are either milked in their tie stalls or are led into a separate milking center. Milking centers (also called parlors) are separate buildings, apart from the lactating cow confinement. Approximately 60 percent of dairy operations reported that they milk the cows from their tie stalls, while 40 percent reported that they used a milking center. However, 78 percent of the lactating cow population is milked in a milking center. Therefore, it can be interpreted that many of the large dairies are using milking centers, while the smaller dairies are typically using tie stalls.

2.1.3 Manure Management

Dairy manure management systems are generally designed based on the physical state of the manure being handled. Dairy cattle manure is collected and managed as a liquid, a semi-solid or slurry, and a solid, and most dairies have both wet and dry manure management systems.

In a slurry or liquid system, manure is flushed from alleys or pits to a storage facility. Typically, effluent from the solids separation system or supernatant from ponds or anaerobic lagoons is used as flush water. The supernatant is the clear liquid overlying the solids that settle below. Dairy manure that is handled and stored as a slurry or liquid may be mixed with dry manure. Liquid systems are usually favored by large dairies for their lower labor cost and because the larger dairies tend to use automatic flushing systems.

Manure accumulates in confinement areas such as barns, drylots, and milking center, and is primarily deposited in areas where the herd is fed and watered. Drylots are used to house calves, and heifers. Due to loss of moisture through evaporation and drainage, drylot manure can either be spread directly after collection or stored in stockpiles for subsequent disposal by land application. Either drylots or freestall barns are used to house the lactating herd when they are not milked. Dairy cattle manure accumulations in freestall barns are typically collected and removed by mechanized scraping systems or by using a flush system. The milking center houses the lactating herd when they are being milked.

2.1.3.1 Manure Storage

Solid manure (from the feedlot and from scraped freestall barns) is typically stored in uncovered storage stockpiles. Because open piles are subjected to rain, they exhibit emission profiles of both aerobic and anaerobic conditions over time. When wet, the stockpiles will be potential sources of NH_3, H_2S, nitrous oxide, and odor causing compounds due to anaerobic decomposition. When dry, they will be emission sources of nitrous oxide from aerobic decomposition, and particulate matter.

Manure handled as a slurry or liquid is stored in earthen impoundments (e.g., anaerobic lagoons). Above ground tanks are another option for storage of these types of manures but are not commonly used. Storage impoundments are designed to hold the total volume of manure and process wastewater generated during the storage period, the increase in volume due to normal precipitation and the increase in volume due to the 25-year, 24-hour storm event while maintaining a minimum freeboard depth of one foot at all times. Emissions from storage tanks and ponds include NH_3, H_2S, VOCs and CH_4. The magnitude of emissions will depend primarily on the length of the storage period and temperature of the manure. Low temperatures will inhibit the microbial activity responsible for the creation of H_2S and CH_4, but may increase VOC emissions and odors. Long storage periods will increase the opportunity for emissions of VOCs, H_2S and NH_3.

2.1.3.2 Solids Separation

In the dairy industry, liquid-solids separation may be used to remove solids from run-off collected from drylots and flushed manure from freestall barns and milking centers. The liquid from solids separation is sent to a storage pond or anaerobic lagoon; the solid is stored in piles. Solids separation is necessary to reduce the organic loading to storage ponds and lagoons so they do not overflow. Mechanical separators (stationary screens, vibrating screens, presses, or centrifuges) or gravity settling basins may be used for this purpose. Emissions from separation activities are dependent on how frequently solids are removed. If solids remain in settling basins and mechanical separation systems longer, emissions of NH_3, H_2S, VOCs and CH_4 may be significant. Generally, the retention time of separation activities is short (i.e., less than one day).

2.1.3.1 Waste Stabilization

Stabilization is the treatment of manure to reduce odor and volatile solids prior to land application. Run-off from drylots and liquid manure from flush alleys are often stabilized in anaerobic lagoons. Anaerobic lagoons use bacterial digestion to decompose organic carbon into CH_4, CO_2, water and residual solids. Single-cell systems combine both stabilization and storage in one earthen structure whereas two-cell systems separate stabilization and storage (i.e., anaerobic lagoon followed by a storage pond).

2.2 Overview of Swine Industry

Swine operations are those that deal with the breeding and growth of pigs for meat. Typical swine operations combine various stages of swine development. The number of swine sites in the U.S. 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 hog industry becoming increasingly dominated by large totally enclosed confinement operations capable of handling 5,000 hogs or more at a time.

In 2007, there were 75,442 swine operations in the U.S. These operations produced 89.6 million pigs. Farms vary in size from operations with a few hundred pigs to some newer operations that house hundreds of thousands of animals at one time. These data show the increasing dominance by large operations. In 2006, 88 percent of the farms had a capacity of 2,000 pigs or less. These smaller operations confined 20 percent of the total inventory of pigs. In contrast, larger operations, which represent 12 percent of the number of farms, confined 80 percent of the inventory. The largest 4 percent of farms (>5000 head) confined 54 percent of the inventory. Swine production historically has been centered in the Midwest. As of 2006, Iowa was the largest hog producing state with 17.3 million pigs, followed by North Carolina (9.5) Minnesota (6.9), Illinois (4.2) and Indiana (3.4).

2.2.1 Production Cycle

The production cycle for hogs 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 pigs or 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 of 8 to 10 piglets per litter 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. 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 hogs normally are slaughtered at about 26 weeks of age.

Swine operations can be of several types. According to the 2007 USDA Census of Agriculture, the most common operation is the growing-finishing operation, followed by the farrow-to-finish operation that encompasses all three phases 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.

2.2.2 Animal Confinement

Although there are still many operations where pigs are raised outdoors, the trend in the swine industry is toward larger operations where pigs are raised in totally or partially enclosed confinement facilities. Typically, the gestation and farrowing, nursery, and grow-finish phases of the production cycle occur in separate, specially designed facilities. Farrowing operations require intense management to reduce piglet mortality. Houses will have farrowing pens, and the piglets are provided a protected area of about 8 square feet. 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 diseases from one herd to another. Finishing pigs require less intensive management and can tolerate greater variations in environmental conditions without incurring health problems.

2.2.3 Manure Management

Although use of open lots for swine production still occurs, this method of confinement generally is limited to small operations. Swine manure produced in open lots is handled as a solid in a similar fashion as at beef cattle feedlots and dairy cattle drylots. In enclosed confinement facilities, swine manure is handled as either a slurry or a liquid.

There are four principal types of waste management systems used with total and partially enclosed confinement housing in the swine industry: deep pit, pull-plug pit, pit recharge, and flush systems. 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. These practices do not represent all of the practices in use today; however, they are the predominant practices currently used by swine operations.

2.2.3.1 Manure Storage

Most large hog farms have from 90 to 365 days of manure storage capacity (NPPC, 1996). Storage is in either an anaerobic lagoon or a storage facility. Typical storage facilities 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. Anaerobic lagoons emit less VOCs and noxious odors than storage facilities, but emit more CH_4.

Storage facilities and anaerobic lagoons are operated differently. Storage facilities hold manure until the vessel is full and then are fully emptied at the next available opportunity. To maintain proper microbial balance, lagoons are sized for a design manure acceptance rate and are emptied on a schedule (but are never fully emptied). This section describes the types of lagoons and storage facilities used and the factors affecting their design.

Storage facilities include deep pits (beneath confinement buildings), in-ground tanks, above-ground tanks, and earthen ponds. Most storage facilities are open to the atmosphere. Manure storage tanks and earthen ponds not only must have adequate capacity to store the manure produced during the storage period but must also store any process wastewaters or runoff that require storage. In addition, provision for storage of the volume of settled solids that will accumulate for the period between solids removal is necessary. Due to the size of storage structures for liquid and slurry type manures, it is difficult to completely mix and empty these facilities during draw down at the end of each storage period. Thus, an accumulation of settled solids will occur requiring a complete clean out of the facility periodically. Estimates of rates of settled solids accumulation for various manures can be found in the Agricultural Waste Management Field Handbook (USDA, 1992).

The microbial processes responsible for CH_4 and VOC formation also occur in storage tanks and ponds. However, the necessary balance in microbial populations for the complete reduction of organic carbon to CH_4 and CO_2 is never established due to higher organic loading rates and accumulations of high concentrations of VOCs, which inhibit CH_4 formation. Thus, emissions of CH_4 from manure storage tanks and ponds will be lower than at anaerobic lagoons, and emissions of VOCs will be higher. Rates of formation of NH_3 and H_2S will not differ, but emission rates may differ depending on hydraulic retention time, pH and the area of the liquid-atmosphere interface. The pH of storage facilities normally will be acidic due to the accumulation of organic acid, which will reduce the rate of NH_3 emission but increase the rate of H_2S emission. The reverse is true for anaerobic lagoons, which have pH values that typically are slightly above neutral. However, time and surface area are probably the more significant variables controlling the masses of NH_3 and H_2S emitted.

2.2.3.2 Waste Stabilization

The anaerobic lagoon has emerged as the overwhelmingly predominant method used for the stabilization and storage of liquid swine manure. Methods of aerobic stabilization (e.g., oxidation ditches or aerated lagoons) were abandoned many years ago due to high electricity costs and operational problems such as foaming. Solids separation is typically not practiced at swine facilities.

Several factors have contributed to the use of anaerobic lagoons for swine waste management. One is the ability to handle the manure as a liquid and use irrigation for land application. A second is the potential to reduce noxious odors by maximizing the complete reduction of complex organic compounds to CH_4 and CO_2, which are odorless gases. Finally, the use of anaerobic lagoons in the swine industry was driven, in part, by the potential to maximize nitrogen losses through NH_3 volatilization thereby reducing land requirements for ultimate disposal. With the shift to phosphorus as the basis for determining acceptable land application rates for animal manures, maximizing nitrogen loss is ceasing to be an advantage.

The design and operation of anaerobic lagoons for swine and other animal manure has the objective of maintaining stable populations of the microorganisms responsible for the reduction of complex organic compounds to CH_4 and CO_2. The microbial reduction of complex organic compounds to CH_4 and CO_2 is a two-step process, in which a variety of VOCs are formed as intermediates. Many of these VOCs, such butyric acid, are sources of noxious odors when not reduced further to CH_4. Methanogenic microorganisms have slower growth rates than the microbes responsible for the formation of VOCs. Therefore, anaerobic lagoons must be designed and operated to maintain a balance between the populations of these microorganisms and methanogens to avoid accumulations of VOCs and releases of associated noxious odors.

Both single cell and two cell systems are used for the stabilization and storage of swine manure. In single cell systems, stabilization and storage are combined. In a two-cell system, the first cell has a constant volume and provides stabilization while the second cell provides storage.

With two cell systems, water for pit recharge or flushing is withdrawn from the second cell. In climates with low precipitation and high evaporation rates, there may be one or more additional cells for the ultimate disposal of excess liquid by evaporation. Anaerobic lagoons use bacterial digestion to decompose organic carbon into CH_4, CO_2, water, and residual solids. Periodic removal of settled solids will be necessary. Typically, lagoons are dredged every 10 to 15 years, and the sludge is applied to land.

The design of lagoon treatment cells is similar to storage ponds with one exception. Lagoons are never completely emptied, except when accumulated solids are removed. Lagoons require permanent retention of what is known as the minimum treatment volume that should be reflected in design. Thus, lagoons must be larger in total volume than ponds that provide storage for the same volume of manure.

Determination of minimum treatment volume for lagoons is based on Natural Resources Conservation Services recommended total volatile solids (TVS) loading rates and the daily TVS loading to the lagoon. For anaerobic lagoons, recommended rates range from 3 lb TVS per 1,000 ft^3 per day in northern parts of Montana and North Dakota to 12 lb TVS per 1,000 ft^3 per day in Puerto Rico and Hawaii. This is a reflection of the effect of temperature on the rate of microbial activity. The calculation of minimum treatment volume is simply the daily TVS loading to the lagoon divided by the recommended TVS loading rate for the geographical location of the lagoon (USDA, 1992).

With open manure storage tanks, ponds, and lagoons, provision also is necessary to store the accumulation of normal precipitation directly falling into the structure less evaporation during the storage period. The storage requirement for normal precipitation less evaporation varies geographically. In addition, there are provisions for storage of precipitation from a 25-year, 24-hour storm event, which also varies geographically, with a minimum of one foot of free board remaining. Design values used for the accumulation of normal precipitation less evaporation are based on mean monthly precipitation values for the location of the storage facility obtained from the National Oceanic and Atmospheric Administration.

In some situations, manure storage ponds or lagoons also may be used for the storage of runoff captured from open confinement areas. In these situations, provision for storage of runoff collected from normal precipitation during the storage period as well as from a 25-year, 24-hour storm event must be included in the design storage capacity of the pond. Expected annual and monthly runoff values for the continental U.S., expressed as percentages of normal precipitation, for paved and unpaved open lots can be found in the Agricultural Waste Management Field Handbook (USDA, 1992).

2.3 Emissions from Open Sources

Animal feeding operations emit particulate and gaseous substances during the three primary stages: confinement, manure storage and treatment, and land disposal. For the manure storage and treatment stage, emissions generally include H_2S, NH_3, methane (CH_4), VOCs and carbon dioxide (CO_2).

Open sources primarily produce gaseous emissions that are the products of the microbial decomposition of manures. Decomposition and the formation of these gaseous compounds begin immediately at excretion and will continue until the manure is incorporated into the soil. Therefore, the substances generated and the subsequent rates of emission depend on a number of variables, including the species of animal being produced, feeding practices, type of confinement facility, type of manure management system, and land application practices.

Emissions generation at any point in the process depends on several factors. The potential for PM emissions depends on whether the manure is handled in a wet or dry state. The potential for gaseous emissions generally depends on several factors: (1) the presence of an aerobic or anaerobic microbial environment, (2) the precursors present in the manure (e.g., sulfur), (3) pH of the manure, and (4) time and temperature in storage, which primarily affects mass emitted. To form H_2S (and other reduced sulfur compounds), CH_4, and VOCs requires an anaerobic environment. Therefore, the potential to emit these substances is greatest when manure is handled as a liquid or slurry. Ammonia will be generated in both wet and dry manure. Nitrous oxide will be formed only when manure that is handled in a dry state becomes saturated (thus forming transient anaerobic conditions). Emissions of NH_3 and H_2S are influenced by pH. The manure pH affects the partitioning between these compounds and their ionized forms (NH_4^+ and HS^-), which are nonvolatile.

Temperature has two effects: (1) Temperature affects gas phase vapor pressure, and therefore, the volatility. For substances that are soluble in water (i.e., NH_3, some VOCs, H_2S and other reduced sulfur compounds), emissions will be greater at higher temperatures. Emission rates of these substances will be greater in warmer climates and in the summer rather than winter. Methane is insoluble in water, and at any temperature will be emitted very quickly after formation. (2) Higher temperature favors the microbial processes that generate CH_4 and other substances.

Long periods of manure residence time in, either confinement, storage or stabilization facilities, provide greater opportunities for anaerobic breakdown and volatilization to the air. In addition, masses emitted will increase with time. The amount of sulfur ingested by an animal will affect the potential for H_2S production in manure. Sulfur can be present in feed additives and, in some cases, from water supplies. The amount of nitrogen in feed (proteins and amino acids) affects NH_3 and nitrous oxide emission potential. The amount of carbon affects CH_4 and CO_2 potential. Ensuring that the composition of feedstuffs does not exceed the nutritional needs of the animal will reduce emissions.

2.4 NAEMS Open-Source Monitoring Sites

The following section provides an overview of the NAEMS open source monitoring sites listed in Table 2-1. For more detail on the sites and the monitoring conducted, please see the site monitoring plans and final reports provided in Appendix D.

2.4.1 Dairy Sites

Three dairy lagoons were monitored under the NAEMS (Table 2-1). The sites were selected to capture different stages and manure practices typical of the industry. The sites selected also represent the broad geographical extent of dairy production to also represent different climatological settings for farm and any regional differences in farm practices.

Dairy lagoon emissions were to be measured continuously at one farm (IN5A) for one year and for up to 21 days each season for two years at the three remaining farms (WA5A and WI5A). Sampling periods for each site are also listed in Table 2-1.

Table 2-1. Summary of NAEMS Open-Source Sites

Site	Phase[a]	Source Type	Manure Collection	Monitoring Period									
				1	2	3	4	5	6	7	8	9	10
IN4A	Sow	Lagoon	PP[b]	7/1/07 - 8/31/07	9/1/07 - 11/30/07	12/1/07 – 2/28/08	3/1/08 – 5/31/08	6/1/08 - 7/14/08					
NC4A[c]	Sow	Lagoon	PPR[d]		10/4/07 - 10/22/07	1/29/08 - 2/11/08	3/31/08 - 4/16/08	8/13/08 - 9/2/08	9/4/08 - 9/23/08	1/14/09 - 2/2/09	4/28/09 - 5/11/09	7/1/09 - 7/21/09	
OK4A[c]	Sow	Lagoon	PPR	6/27/07 - 8/29/07	11/7/07 - 11/27/07	11/28/07 - 12/18/07	4/23/08 - 5/6/08		10/1/08 - 10/15/08	1/8/09 - 1/27/09	4/1/09 - 4/21/09	6/25/09 - 7/14/09	
IA3A	Finisher	Lagoon	PP		8/30/07 - 9/26/07	12/19/07 - 1/15/08	5/14/08 - 6/4/08	6/4/08 - 6/25/08	11/13/08 - 11/25/08	11/25/08 - 12/16/08	4/8/09 - 4/23/09	7/28/09 - 8/17/09	
NC3A	Finisher	Lagoon	PPR		10/24/07 - 11/7/07	2/13/08 - 3/5/08	3/6/08 - 3/26/08		9/25/08 - 10/14/08	2/4/09 - 2/23/09	5/12/09 - 6/2/09	6/2/09 - 6/22/09	9/24/09 - 12/1/09
OK3A	Finisher	Lagoon	PPR		8/30/07 - 9/18/07	1/24/08 - 2/19/08	5/7/08 - 5/29/08	5/29/08 - 6/10/08	11/5/08 - 12/2/08	12/2/08 - 12/16/08	4/23/09 - 5/14/09	7/15/09 - 8/4/09	
IN5A	Dairy	Lagoon	Scrape						9/11/08 - 11/30/08	12/1/08 - 2/28/09	3/1/09 - 5/31/09	6/1/09 - 8/17/09	
WA5A	Dairy	Lagoon	Flush			2/25/08 - 3/12/08	3/12/08 - 3/26/08	8/8/08 - 9/3/08	9/3/08 - 9/26/08		5/18/09 - 6/4/09	6/4/09 - 6/20/09	
WI5A[c]	Dairy	Lagoon	Flush	7/18/07 - 8/28/07	11/13/07 - 11/28/07	11/28/07 - 12/18/07	4/23/08 - 5/13/08	6/25/08 - 7/14/08	10/21/08 - 11/11/08	12/17/08 - 1/7/09	3/10/09 - 4/7/09		

[a] Characterizes type of farm.
[b] PP= pull-plug pit.
[c] Area site that also had barns sites.
[d] PPR = pull-plug pit with recharge.

2.4.1.1 IN5A

The Indiana dairy consisted of three barns, a feed storage area, special needs barn, milking parlor, an office and tool and repair shops (Figure 2-1). Construction of the dairy was completed in 2002. The farm has a capacity of 2,600 cows.

Manure was vacuumed from the lactating cow barns and special needs barn every 12 hrs and placed in lagoons near the barns. Manure was flushed from the holding area and milking parlor every half hour. Manure from the barns was scraped into pits that are located at the end of each barn (Figure 2-1). A small fraction of scraped waste was held in a slurry tank. The wastewater (flush) from the holding area and milking parlor was transferred into a rectangular settling lagoon south of the road then into the waste lagoon south of the road. The inlet to the lagoon was located at its north end. The clay-lined waste lagoon was 85 m (280 ft) wide and 116 m (380 ft) long, and is oriented north to south. At maximum capacity, the liquid depth was 5 m (16 ft) with a surface area of 9,884 m^2 (106,400 ft^2) and volume of 48,200 m^3 (1,702,400 ft^3). Sludge had never been removed from the lagoon.

Figure 2-1. Aerial View of IN5A

2.4.1.2 WA5A

The Washington farm consisted of six barns, a milking parlor, and an office (Figure 2-2). The facility had a capacity of 4,400 milking cows and 1,200 dry cows in three units. Construction of the dairy was completed in 2002.

The farm had freestall style barns, with automated flushing that occurred four times daily. Manure was transferred to the upper lagoon from a sand separation pit. The two earthen-lined settling lagoons were located to the south of the barns, and were used in alternate years. Liquids were skimmed, separated and returned as flush to the barns. One lagoon was actively filled while the other was drying or sludge was being entirely removed. The settled solids (sludge) were completely removed within a year by front-end loader. These removed solids were then strained through screens and centrifugal/screw presses with the liquid transferred to large serpentine concrete basins for secondary settling. These settled solids were then dried for bedding. Removed water from the settled solids was stored in a large clarified water storage basin for dilution of barn flush water from the lagoon. Sludge from the settling basins prior to complete removal was periodically applied on surrounding land utilizing underground pressurized pipes and a no-till soil injector. Wastewater entered the measured lagoon from the northwest corner.

Gaseous emissions occur both during lagoon filling and during sludge removal. The east lagoon was rectangular with dimensions of 183 m (600 ft) by 72 m (235 ft). The west lagoon was five-sided with dimensions of approximately 183 m (600 ft) long and 83 m (271 ft) wide with the southwest corner of the lagoon cut off. The east lagoon was measured for gaseous emissions. At maximum capacity this lagoon had a liquid depth of 5 m (18 ft), surface area of 13,098 m^2 (141,000 ft^2) and a volume of 186,300 m^3 (2,005,500 ft^3). Sludge was last removed from the lagoon in 2006.

Figure 2-2. Aerial View of WA5A

2.4.1.3 WI5A

The Wisconsin farm had a total of six barns, a milking parlor with holding pen and a special needs area. The farm had a capacity of 1,700 Holstein cows. Construction of the dairy was completed in 1994.

Manure from the freestall barns and the milking parlor complex was removed by flushing three times daily. The manure flushed from the parlor, holding pen, and freestall barns flowed to a solids separator, from which the solids are removed and stacked on a pad until they are spread on fields (the nearest of which is approximately 100 m (328 ft) from the barns) (Figure 2-3). The liquid effluent from the solids separator was pumped back into vertical tanks for reuse to flush the barns. Once a week, enough water was removed from the third stage of the three-stage lagoon and added to the flush tanks to compensate for water lost in the recycled flush system. The lagoons were pumped out into trucks twice yearly. The first and second stages of the three-stage lagoon system were monitored (Figure 2-3). The first lagoon had a width of 52 m (170 ft) and length of 82 m (270 ft). At maximum capacity, the first lagoon had a surface area of 4,264 m^2 (45,900 ft^2) and a volume of 10,561 m^3 (373,000 ft^3). The second lagoon had a width of 37 m (120 ft) and length of 79 m (260 ft). At maximum capacity, the second lagoon had a surface area of 2,898 m^2 (31,200 ft^2) and a volume of 6,420 m^3 (226,700 ft^3). Both lagoons had liquid depths of 3 m (11 ft) and sludge was last removed from the second lagoon in 2006.

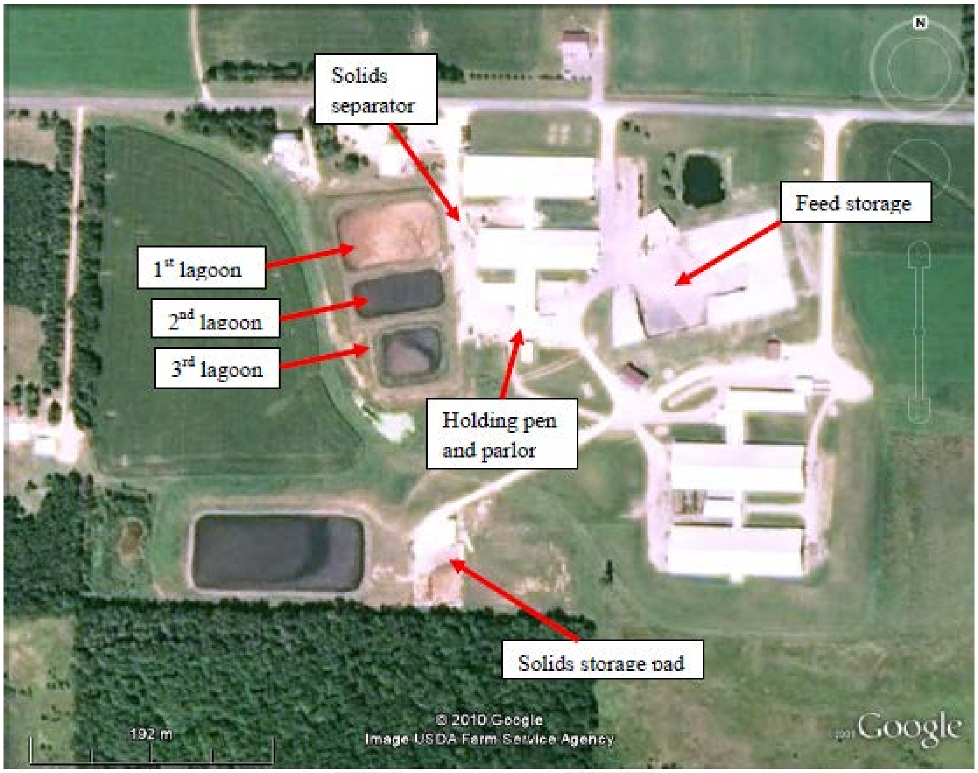


Figure 2-3. Aerial View of WI5A

2.4.2 Swine Sites

The six swine farms that had lagoons monitored as part of the NAEMS are listed in Table 2-1. The swine manure lagoon emissions were measured continuously at one farm (IN4A) for one year and up to 21 days for each season over two years at the remaining farms (IA3A, NC3A, NC4A, OK3A, OK4A). The sampling periods for each site are also listed in Table 2-1. Sites for monitoring were selected to capture different stages (i.e., breeding-gestation and growing-to-fishing) and manure practices typical of the industry. The sites also represent the broad geographical extent of swine production to also represent different climatological settings for farm and any regional differences in farm practices.

2.4.2.1 IN4A

The Indiana breeding/gestation farm consisted of nine barns and a lagoon (Figure 2-4), and had a capacity of 1,400 sows. The facility had been added to for many years, starting operations in 1968, while the last building addition was completed in 1992. In 1998, the facility was changed from a finisher operation to a farrow-to-wean operation.

Liquid waste from the deep pits of the barns was transferred once every two weeks to the lagoon by a single inlet on the east side of the lagoon (Figure 2-4). The lagoon was south from the barns. The clay-lined waste lagoon was 111 m (365 ft) by 115 m (378 ft). At maximum capacity, the liquid depth was 4 m (12 ft) with surface area of 11,240 m^2 (121,000 ft^2) and a volume of 33,975 m^3 (1,200,000 ft^3). Sludge has never been removed from the lagoon. During the growing season, corn completely surrounded the lagoon.

Figure 2-4. Aerial View of IN4A

2.4.2.2 NC4A

The breeding/gestation farm in North Carolina consisted of three barns, one each of gestation, breeding, and farrowing, and an office (Figure 2-5). The facility had a capacity of 2,000 sows in three units. Construction of the farm was completed in 1994.

Manure from the barns was transferred once a week from the gestation, farrowing, and breeding barns to the lagoon by pull plug with lagoon water recharge of the pits. Wastewater from all three buildings combined into one inlet (SW corner of lagoon- Figure 2-5). The waste lagoon was located to the north of the barns. The clay-lined, trapezoidal-shaped lagoon was oriented east to west and measured 125 m (410 ft) wide and 190 m (624 ft) long. The lagoon had a surface area of 23,193 m^2 (249,672 ft^2) and a volume of 56,851 m^3 (2,077,450 ft^3). At the beginning of the NAEMS, the sludge depth was approximately 0.7 m (2 ft). Liquid was removed as weather permitted. Sludge from the lagoon had not been removed since construction (15-yr sludge removal cycle). Barns on this farm were also monitored as a part of the NAEMS.

Figure 2-5. Aerial View of NC4A

2.4.2.3 OK4A

The Oklahoma breeding/gestation farm consists of three barns and one office (Figure 2-6). The facility has a capacity of 2,784 sows. Construction of the sow farm was completed in 1994.

Manure from the barns was transferred weekly from the two gestation units and every 2.5 weeks from the farrowing unit to the lagoon by pull plug and lagoon water recharge. Waste water from the two gestation units combined into one inlet (the southerly inlet in Figure 2-6), while wastewater from the farrowing unit entered the lagoon from the northerly inlet (Figure 2-6). The rectangular waste lagoon was located to the east and was separated by a drainage swale from the barns. The clay-lined lagoon was 117 m (383 ft) wide and 193 m (632 ft) long, and was oriented north to south. Liquid depth was approximately 5.5 m (18 ft). The lagoon had a surface area of 22,486 m^2 (242,056 ft^2) and volume was approximately 72,800 m^3 (4,357,008 ft^3). Sludge from the lagoon has not been removed since construction (20-yr sludge removal cycle). Field applications occurred up to two times per year, based on rainfall.

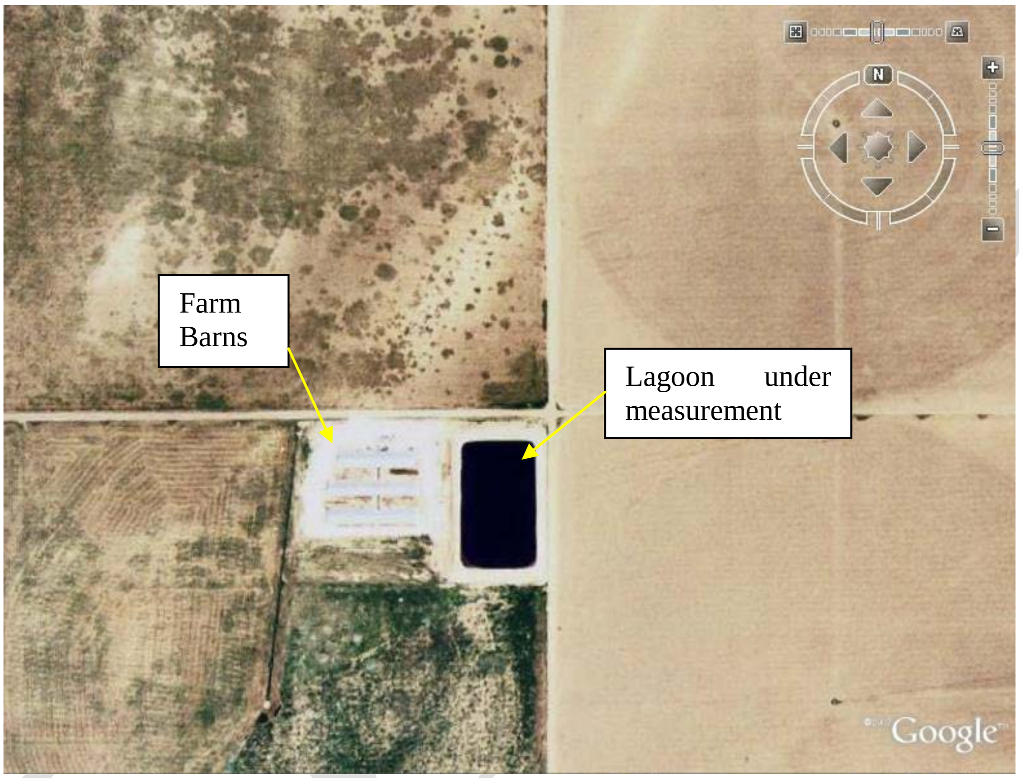


Figure 2-6. Aerial View of OK4A

2.4.2.4 IA3A

The growing/finishing farm in Iowa consisted of four barns and a manure lagoon (Figure 2-7). The facility had a capacity of 3,840 finishers in the four units. The construction of the facility was completed in 1998.

Manure from the 2 ft deep pits in each of the four barns was transferred to the lagoon, which was west of the barns (Figure 2-7), approximately once every ten weeks through two inlets (Figure 2-7). The concrete, circular structure had a diameter of 55 m (180 ft) with its sides approximately 0.5 m (1.5 ft) above and 2 m (6.5 ft) below ground level. At maximum capacity the structure had a liquid depth of 2 m (7 ft), surface area of 2,363 m^2 (25,442 ft^2) and a volume of 5,763 m^3 (203,537 ft^3). Sludge had never been removed from the lagoon.

Figure 2-7. Aerial View of IA3A

2.4.2.5 NC3A

The North Carolina growing/finishing farm consisted of five barns (Figure 2-8) and an office, in addition to the lagoon itself. The facility had a capacity of 8,000 finishing pigs in five units. Construction of the farm was completed in 1996.

Manure from the barns was transferred daily to the lagoon by pull plug and lagoon water recharge. Wastewater combined into one inlet (Figure 2-8). The rectangular waste lagoon was located to the north and was separated by a drainage swale from the barns. The clay-lined lagoon was 113 m (370 ft) wide and 173 m (568 ft) long and was oriented east to west. The lagoon had a maximum liquid depth of 3.3 m (11 ft), a surface area of 18,986 m^2 (204,386 ft^2) and a volume of 45,961 m^3 (1,623,326 ft^3). Wastewater was removed for irrigation as weather permitted. Sludge from the lagoon had not been removed since construction (15-year sludge removal cycle).

Figure 2-8. Aerial View of NC3A

2.4.2.6 OK3A

The Oklahoma growing/finishing farm consisted of three barns (Figure 2-9). The facility had a maximum capacity of 3,024 finishing pigs. Construction was completed in 1997.

Manure from the barns was transferred three times a week to the lagoon by a pull plug system with lagoon water recharge. Wastewater from all three units combined into one inlet. The waste lagoon was rectangular and was located to the west of the barns (separated by a drainage swale). The clay-lined lagoon was 59 m (193 ft) wide and 208 m (683 ft) long, and was oriented north to south. At maximum capacity, the lagoon had a liquid depth of 6 m (20 ft) with a surface area of 11,203 m^2 (120,600 ft^2) and a volume of 28,637 m^3 (1,011,463 ft^3). Liquid was removed approximately every six months. Sludge from the lagoon had not been removed since construction (20-yr sludge removal cycle).

Figure 2-9. Aerial View of OK3A

3.0 DATA AVAILABLE FOR EEM DEVELOPMENT

In the Air Compliance Agreement, the EPA committed to developing EEMs for estimating daily and annual emissions from swine and dairy lagoons using the emissions and process data collected under the NAEMS and any other relevant data and information that are available. Section 3.1 summarizes the NAEMS emissions and process data for swine and dairy lagoons. Section 3.2 discusses the other relevant data that the EPA has gathered both under a Call for Information (CFI) that was issued by the Agency on January 19, 2011 and through previously-conducted literature searches.

3.1 NAEMS Data

3.1.1 Data Received

The EPA received final reports and data spreadsheets for three swine breeding and gestation operations (sites IN4A, NC4A and OK4A), three swine growing and finishing operations (sites IA3A, NC3A and OK3A), and three dairy operations (sites IN5A, WA5A and WI5A). In general, the final reports for each site describe the monitoring locations and sampling methods and present the results of the emissions measurements expressed in various units (e.g., daily average emissions in kg NH_3/day, g NH_3/day-head and g NH_3/day-animal unit). The final reports also contain summaries of producer activities at the site (e.g., lagoon pump-out events), notable weather conditions, lagoon appearance and the results of liquid sampling events that occurred during the monitoring period. Appendix D contains the final reports submitted for each monitoring site. Microsoft Excel® spreadsheets containing 30-minute values for emissions and lagoon condition data and 5-minute values for meteorological data were also submitted to the EPA.

To increase public involvement and maintain transparency throughout the EEMs development process, the EPA has made information and data relating to the NAEMS available at http://www.epa.gov/airquality/agmonitoring/. This website provides links to background information regarding the Air Compliance Agreement, the NAEMS (including information describing the monitoring sites, site-specific data files and final reports) and the EPA's CFI. Additionally, the EPA has included all information received pertaining to the NAEMS in the public docket (EPA Docket ID No. EPA-HQ-OAR-2010-0960), which is available at: http://www.regulations.gov.

Table 3-1 summarizes the emissions and process data elements that were required to be monitored by either the NAEMS Monitoring Protocol, the QAPPs, the SMPs or the SOP documents submitted to the EPA. The NAEMS Monitoring Protocol was developed in a collaborative effort by representatives from the EPA, USDA, AFO industry representatives, agricultural researchers, state and local air quality agencies and environmental organizations. The NAEMS Monitoring Protocol identified the parameters to be monitored during the study and, for some parameters, specific measurement methodologies and frequencies. For those parameters for which either or both the measurement methodology and frequency was not specified, the study's Science Advisor provided the specific information in the study's QAPP. Table 3-1 also provides specific information regarding data availability that is based upon the EPA's review of the final reports and data spreadsheets.

During its review of the final reports and data spreadsheets, the EPA identified missing emissions and process data. A summary of the issues and discrepancies identified by the EPA's review are presented in Section 5. The EPA's review of the data spreadsheets also identified that the daily and 30-minute emissions values reported by the NAEMS researchers contained negative values. After discussion with the study's Science Advisor, it was determined the negative values were a result of instrumentation drift, and are considered to be valid values. To avoid possible complications with EEM development (e.g., the EEM predicting negative emissions), the negative values were withheld from the data sets used for EEM development. The amount of measured negative values is low (less than 2 percent) compared to the total number of emissions records for NH_3, which indicates that the steps taken to calibrate and maintain instrumentation and to minimize the influence of other on-site sources of ambient NH_3 emissions were reasonably effective. Because of their relatively small number, excluding the negative values does not compromise the EEM datasets for NH_3. The EPA's review of the data spreadsheets also identified daily average emissions and parameter values that were exactly zero. After further review and discussion with the study's Science Advisor, it was determined the values were valid reported values and were used for EEM development.

Table 3-1. NAEMS Emissions and Process Parameter Data Received

Parameter Information				NAEMS Data		
Parameter	**Required by the NAEMS Monitoring Protocol**	**Measurement Methodology**	**Measurement Frequency**	**Data Description**	**Measurement Methodology**	**Measurement Frequency**
Concentrations						
NH_3	Yes	Fourier transform infrared spectroscopy (FTIR) and UV differential optical absorption spectroscopy (UV–DOAS)	Not specified	Average measured concentration values	TDLAS/ Boreal Laser, Inc. GasFinder2, INNOVA 1412 multi-gas analyzer	Logged every 15 & 60 seconds, for up to 21 days per season; 30-minute averages and daily averages provided
H_2S	Yes	UV differential optical absorption spectroscopy (UV–DOAS)	Not specified	Average measured concentration values	UV differential optical absorption spectroscopy (UV–DOAS), pulsed fluorescence	Logged every 15 & 60 seconds, for up to 21 days per season; 30-minute averages and daily averages provided
Hydrocarbons emissions (continuous)	Yes	Fourier transform infrared spectroscopy (FTIR)	Not specified	Data not received [a]	PAS/ INNOVA 1412 multigas analyzer	Logged every 15 & 60 seconds, for 24 months
VOC Samples	Yes	Analysis of liquid samples	Not specified	Data not collected [b]		
Emissions						
NH_3	Yes	backward Lagrangian stochastic method (bLS) computed tomography (CT) method using Eulerian Gaussian statistics and a fitted wind profile from the two 3D sonic anemometers	Not specified	Average calculated emission rate values	Radial Plume Mapping (RPM) & backward Lagrangian stochastic method (bLS)	30-minute averages and daily averages provided
H_2S	Yes			Average calculated emission rate values	backward Lagrangian stochastic method (bLS) & Ratiometric Method	30-minute averages and daily averages provided
Hydrocarbons emissions (continuous)	Yes			Data not received [c]	Ratiometric to RPM Model	Data not received [c]
VOC Samples	Yes	WATER9 model	Not specified	Data not collected [b]		
Lagoon Parameters						
Lagoon temperature	Yes	Not specified		5-minute average values	Thermistor	Logged every minute for up to 21 days per season; 5-miute average values provided
pH	Yes	Not specified		5-minute average values	Electrochemical pH meter	Logged every minute for up to 21 days per season; 5-miute average values provided
Oxidation Reduction Potential	No	Not applicable		5-minute average values	ORP Sensor	Logged every minute for up to 21 days per season; 5-miute average values provided
Lagoon Solids depth	No	Not applicable		Partial data received [c]	Sludge level detector	One measurement per period at IN4A, NC4A, OK4A, NC3A and OK3A
Ambient Meteorological Parameters						
Wind speed	Yes	3D sonic anemometers at heights of 2 and 12 meters (m)	Not specified	5-minute average values	3D Sonic anemometer & Cup anemometer	Logged every 15 & 60 seconds for up to 21 day per season; 30-minute averages provided
Wind direction	Yes	3D sonic anemometers at heights of 2 and 12 meters (m)	Not specified	5-minute average values	3D Sonic anemometer & Cup anemometer	Logged every 15 & 60 seconds for up to 21 day per season; 30-minute averages provided
Solar radiation	Yes	Not specified		5-minute average values	Pyranometer	Logged every 15 & 60 seconds for up to 21 day per season; 5-minute averages provided

Table 3-1. NAEMS Emissions and Process Parameter Data Received

Parameter Information				NAEMS Data		
Parameter	**Required by the NAEMS Monitoring Protocol**	**Measurement Methodology**	**Measurement Frequency**	**Data Description**	**Measurement Methodology**	**Measurement Frequency**
Ambient pressure	Yes	Not specified		5-minute average values	Aneroid barometer	Logged every 15 & 60 seconds for up to 21 day per season; 5-minute averages provided
Ambient temperature at site	Yes	Not specified		5-minute average values	Thermistor	Logged every 15 & 60 seconds for up to 21 day per season; 5-minute averages provided
Ambient relative humidity at site	Yes	Not specified		5-minute average values	Hygrometer	Logged every 15 & 60 seconds for up to 21 day per season; 5-minute averages provided
Surface Wetness	No	Not specified		5-minute average values	VAC resistance grid	Logged every 15 & 60 seconds for up to 21 day per season; 5-minute averages provided
Nitrogen Mass Balance						
pH of manure	No	Not applicable		Partial data received [d]	Electrochemical pH meter	Sporadic measurements received for NC4A, OK4A, IA3A, NC3A and OK3A.
NH_4 content of manure	No	Not applicable		Partial data received [d]	Kjeldahl/titrimetric	Sporadic measurements received for NC4A, OK4A, IA3A, NC3A and OK3A.
Solids content of manure	No	Not applicable		Partial data received [d]	Gravimetric	Sporadic measurements received for NC4A, OK4A, IA3A, NC3A and OK3A.
TKN content of manure	No	Not applicable		Partial data received [d]	Kjeldahl/titrimetric	Sporadic measurements received for NC4A, OK4A, IA3A, NC3A and OK3A.

[a] The QAPP mentions this parameter was measured; however, these data are not provided in the final report(s).
[b] Data not expected. This parameter is not referenced in the QAPP or final report.
[c] The QAPP mentions this parameter was to be measured once per period; however, these data are only provided for IN4A, NC4A, OK4A, NC3A and OK3A in the final report(s).
[d] The QAPP mentions this parameter was to be measured once per period or every 90 days; however, these data are only provided for NC4A, OK4A, IA3A, NC3A and OK3A in the final report(s) and do not always fall within monitoring periods. Also, the report is unclear if the range is for nitrogen or TKN.

3.1.2 Emissions Levels Reported in the NAEMS Final Reports

Table 3-2 summarizes the average and maximum emissions cited in the final reports submitted to the EPA. The average and maximum daily values were reportedly based on all valid days over the site monitoring periods. A valid monitoring day is one in which 75 percent of the hourly average data values used to calculate the daily value were valid measurements. An hourly average is considered valid if 75 percent of the data recorded during that hour were valid. Data were invalidated due to special events (e.g., audits, calibrations and maintenance), failure of quality control limits (e.g., unreasonably low or high compared with normal ranges combined with supporting evidence that the values are not correct) or when a sample is contaminated. A summary of the major data invalidation events identified by the NAEMS researchers is provided in the final site reports (see Appendix D). The results of EPA's assessment of data completeness are presented in Section 4.

Table 3-2. Reported Emission Rates for NAEMS Lagoon Sites

Site	Average Farm Capacity[a] (head)	Average Animal Weight[a] (lbs)	Average Daily Emissions (kg/day)[b]		Maximum Daily Emissions (kg/day)[b]	
			NH_3 (RPM)	H_2S (bLS)	NH_3 (RPM)	H_2S (bLS)
IN4A	1,400 (sows)	475 (sows)/ 8 (piglets)	58.4 [c]	0.4	58.4 [c]	2.2
NC4A	2,000 (sows)	433 (sows)	70.8	0.3	102.8	1.0
OK4A	2,784	490 (sows)	177.2	8.4	317.9	25.3
IA3A	3,840	150	47.1	2.2	52.8	12.3
NC3A	8,000	135	58.2	2	117.8	10.4
OK3A	3,024	170	102.8	6.3	184.4	25.4
IN5A	2,600	1,500	23.4	0	34.6	42.9
WA5A	5,600	1,400	NA[d]	NA[d]	NA[d]	NA[d]
WI5A	1,700	1,400	NA[d]	0.6	NA[d]	1.9

[a] The average farm capacity and average animal weight values were based on the values provided by the growers in the SMPs.
[b] The emissions as provided in the final reports.
[c] Only one day met the 75% completeness criteria during the study; therefore the average and maximum daily emission values are the same.
[d] The site did not meet the 75% completeness criteria for any day during the study; therefore there are no average or maximum daily emission values.

3.2 Other Relevant Data

Since 2001, the EPA conducted several literature searches and a CFI to identify data and information that were relevant to support a preliminary investigation into air pollution from large AFOs (see the EPA's *Emissions From Animal Feeding Operations* (draft, August 15, 2001)). The EPA evaluated all of the articles and publications received through its own literature searches and obtained through the CFI to identify data and information that could be useful in developing EEMs for swine and dairy lagoons. In conducting this evaluation, the EPA retained for further consideration those resources that satisfied each of the following conditions:

- The resource pertained to monitoring conducted on lagoons at commercial sites.
- The resource contained emissions rates (e.g., mass/time, mass/animal) for NH_3, H_2S or VOC, or data to characterize the inputs or outputs necessary to construct a nitrogen mass balance across the lagoon.
- The resource used methods to measure the emissions concentrations, estimate the emissions rate and characterize mass balance parameters that were consistent with the NAEMS procedures.

The EPA then evaluated the resources that satisfied the Agency's initial review to determine if the data were appropriate for consideration in either developing the EEMs or assessing the predictive accuracy of the EEMs. Section 3.2.1 summarizes the EPA's CFI and the review of the resources obtained. Section 3.2.2 summarizes the Agency's review of the resources obtained by previous EPA literature searches.

3.2.1 CFI

The EPA issued a CFI on January 19, 2011, seeking peer-reviewed, quality-assured emissions and process data relevant to developing EEMs for animal feeding operations. The CFI was designed to help ensure that the Agency would obtain the broadest range of scientific data available. All data and information received by the EPA is contained in the public docket for the Air Compliance Agreement (EPA Docket ID No. EPA-HQ-OAR-2010-0960) and is available online at http://www.regulations.gov.

In the CFI, the EPA requested emissions and process data for AFOs in the broiler, swine, egg-layer, dairy, beef and turkey industries. Although the EPA is interested in all air pollutants emitted from animal confinement, manure storage and treatment and manure land applications sites associated with AFOs, the CFI specifically requested emissions data and related process information for NH_3, H_2S, PM_{10}, $PM_{2.5}$, TSP and VOC.

To ensure compatibility with the NAEMS data, the CFI requested that, to the extent possible, the emissions and related process data provided to the EPA be accompanied by documentation that addresses the following parameters:

General information:

- Description of AFO process measured (e.g., anaerobic lagoon).
- Location of AFO process measured (e.g., physical address, latitude/longitude coordinates of facility).
- Beginning and ending dates of the monitoring period.

Monitoring data:

- Plan for quality assurance and quality control procedures.

- Site-specific monitoring plan.
- Test methods, instrumentation and SOPs used to collect emissions and process data measurements.
- Results of audits conducted on instruments and procedures.
- Field notes and associated documentation collected during the study.
- Emissions data (raw or analyzed) and associated process data.
- Meteorological data, including average ambient temperature, relative humidity, pressure, wind speed, wind direction and insolation (solar radiation) for each day that the study was conducted.
- Sampling dates and times.
- Production data (e.g., quantity of milk produced per day or number of swine produced).
- Calculations and assumptions used to convert concentration data (e.g., ppmv) into mass emissions (e.g., lb/hr).

Animal confinement structures:

- Dimensions of structures monitored.
- Designed and permitted animal capacity.
- Type, age, number and weight of animals contained in the confinement structure over the duration of the monitoring period.
- Manure management system (e.g., pull-plug pit, scrape).
- Manure removal activities over the duration of the monitoring period.
- Ventilation method (i.e., natural or mechanical).
- Calculations and assumptions used to estimate the ventilation rate of the monitored confinement structure.
- Calibration procedures for instruments (e.g., flow meters, fan relays) used to collect data for calculating ventilation rate of the monitored confinement structure.
- Nitrogen content of process inputs and outputs (e.g., feed, water, bedding, eggs, milk).
- Nitrogen content of manure excreted.
- Description of any control device or work practice used in the monitored structure to reduce emissions.

Manure storage and treatment processes:

- Type, age, number, and weight of animals contributing manure to the storage and treatment process over the monitoring period.
- Dimensions of storage/treatment unit monitored (e.g., storage pile, tank, lagoon).

- Design specifications of storage/treatment unit monitored (e.g., volume, volatile solids loading rate).
- Depth of settled solids in storage/treatment unit.
- Temperature, pH and reduction/oxidation potential of manure contained in the storage/treatment unit.
- Moisture, total solids, volatile solids, total Kjeldahl nitrogen and ammoniacal nitrogen content and pH of manure entering storage and treatment process over the monitoring period.

Manure land application sites:

- Type, age, number, and weight of animals contributing manure to the land application site.
- Method used to apply manure (e.g., direct injection, broadcast spreading and frequency of application).
- Area (e.g., acres, square feet) used for manure application over the monitoring period.
- Quantity and moisture content of manure applied.

Table 3-3 and Table 3-4 lists the articles and publications received by the EPA in response to the CFI that pertained to swine and dairy operations, respectively, and their possible application for the NAEMS. As shown in the tables, most of the articles and publications submitted to the EPA did not contain emissions or process data that met the EPA's initial review criteria (e.g., the measurement methods differed from the NAEMS methods). However, a few resources contained material composition data that could be used to supplement the nitrogen mass balance data collected by the NAEMS.

Table 3-3. Review of Swine Lagoon Articles Received in Response to EPA's CFI

Year	Title - Author(s)	Brief Description	Relevant Data	Methodology	Possible Application for NAEMS
2008	*Characterizing ammonia and hydrogen sulfide emissions from a swine waste treatment lagoon in North Carolina* (Blunden, Aneja)	This study investigated the seasonal variations in NH_3 and H_2S flux from a swine waste storage lagoon. The lagoon physicochemical properties were also monitored	NH_3 emissions	NH_3 flux was measured with a dynamic flow-through chamber and a chemiluminescence analyzer (Model 17C).	This study used flux chambers to measure lagoon emissions rather than remote sensing techniques. Thus, applicability to EEM development is limited.
			H_2S emissions	H_2S flux was measured with a dynamic flow-through chamber system and a pulsed fluorescence H_2S/SO_2 analyzer (Model 450C).	
			Lagoon surface water temperature	Two Campbell Scientific 107 temperature probes	
			Lagoon pH	Model CSIM11 pH probe	
			Lagoon water total Kjeldahl nitrogen (TKN)	Samples were submitted to the North Carolina Division of Water Quality (NC DWQ) for TKN analysis.	
			Lagoon water total ammoniacal nitrogen	Samples were submitted to the NC DWQ for total ammoniacal nitrogen (TAN) analysis.	
			Lagoon water total sulfide concentration	Samples were submitted to the NC DWQ for total sulfide analysis.	
			Ambient meteorology: wind speed, wind direction, air temperature, relative humidity, solar radiation	Met One Instruments Model 034-B Windset, Model CS500-L Vaisala 50Y temperature and RH probe, Model LI200X Silicon Pyranometer, Model CSIM11 pH probe	
2000	*Characterization of Atmospheric Ammonia Emissions from Swine Waste Storage and Treatment Lagoons* (Aneja, Chauhan, Walker)	This study investigated the seasonal variations in NH_3 flux from a swine waste storage lagoon.	NH_3 emissions	NH_3 flux was measured using a dynamic flow-through chamber system floating and a series of gas analyzers. NH_3 concentration was determined by subtracting the NOx from the total N.	This study used flux chambers to measure lagoon emissions rather than remote sensing techniques. Thus, applicability to EEM development is limited.
			Lagoon surface water temperature	Continuous measurement with temperature probe (Fascinating Electronics)	

Table 3-3. Review of Swine Lagoon Articles Received in Response to EPA's CFI

Year	Title - Author(s)	Brief Description	Relevant Data	Methodology	Possible Application for NAEMS
			Lagoon pH	Measured continuously using a double junction submersible electrode (Cole Parmer)	
			TKN	Lagoon water samples taken daily and measured for TKN	
2008	*Characterizing Ammonia Emissions from Swine Farms in Eastern North Carolina: Part 1—Conventional Lagoon and Spray Technology for Waste Treatment* (Aneja, Pal Arya, Kim, Rumsey, Arkinson, Semunegus, Bajwa, Dickey, Stefanski, Mottus, Robarge, Williams)	NH_3 flux from waste treatment lagoons and barns at two conventional swine farms were measured. The waste treatment lagoon data were analyzed to elucidate the temporal (seasonal and diurnal) variability and to derive regression relationships between NH_3 flux and lagoon temperature, pH and ammonium content of the lagoon, and the most relevant meteorological parameters.	Lagoon NH_3 emissions	NH_3 flux from the lagoons was measured by a dynamic flow-through flux chamber system and a chemiluminescence analyzer (TEI Model 17C).	This study used flux chambers to measure lagoon emissions rather than remote sensing techniques. Thus, applicability to EEM development is limited.
			Lagoon surface water temperature	Temperature probes (Campbell Scientific Instruments Model 107)	
			Lagoon pH	pH probe (Innovative Sensors)	
			Lagoon water TKN	Not specified.	
			Lagoon water total nitrate and ammoniacal nitrogen	Not specified.	
			Meteorology: wind speed, wind direction, air temperature, relative humidity, solar radiation	Campbell Scientific Instruments CS500 temperature and relative humidity (RH) probe, a LI-Cor 200SZ pyranometer, and a Met One 034A-LC Windset anemometer and wind vane.	
2008	*Characterizing Ammonia Emissions from Swine Farms in Eastern North Carolina: Part 2—Potential Environmentally Superior Technologies for Waste Treatment* (Aneja, Pal Arya, Kim, Rumsey, Arkinson, Semunegus, Bajwa, Dickey, Stefanski, Mottus, Robarge, Williams)	NH_3 fluxes from waste water holding structures and barns at six swine farms using non-lagoon ESTs for waste treatment were measured two two-week periods in warm and cold seasons. NH_3 fluxes were measured on the water-holding structures by a flow through dynamic chamber system. Open-path Fourier transform infrared (FTIR) spectrometers were used	NH_3 flux from surface of water-holding structures	Continuous measurement, recorded 15 minute average, measured with dynamic flow-through chamber system	This study did not examine emissions from conventional swine waste lagoons, but it did compare emissions from these six farms to emissions from a conventional farm lagoon. Because this study used flux chambers to measure emissions, applicability to EEM development is limited.
			NH_3 emissions from the ventilation systems at animal houses	Measured with open-path Fourier transform infrared (OP-FTIR) spectroscopy. NH_3 emissions from barn houses were estimated from average NH_3 concentration measured by OP-FTIR and the rated flow rate for the fan size and setting.	

Table 3-3. Review of Swine Lagoon Articles Received in Response to EPA's CFI

Year	Title - Author(s)	Brief Description	Relevant Data	Methodology	Possible Application for NAEMS
		to measure NH_3 concentrations at the barn vent inlets and outlets for estimating NH_3 emissions from the animal housing units (barns).	Number of pigs, average and total pig mass (kg)	Not specified	
			Feed consumed (kg/pig/week)	Not specified	
			Feed N content (%)	Not specified	
			Nitrogen excretion (kg N/week/1000 kg live weight animals)	N excretion rate (E) in units of kg-N/week/1000 kg-lw) = Fc x Nf x (1-er)/w x 1000, where Fc is the feed consumed (kg/pig/week), Nf is the fraction of N content in feed, er is the feed efficiency rate (ratio of average gain of N to N intake), and w is the average live animal mass (kg/pig).	
2008	*Modeling Studies of Ammonia Dispersion and Dry Deposition at Some Hog Farms in North Carolina (*Bajwa, Pal Arya, Aneja)	This study investigated the NH_3 emissions and local deposition of NH_3 from the lagoons and barns of one hog farm. Dry deposition velocity, dispersion, and dry deposition of NH_3 were studied over different seasons and under different stability conditions using the short-range dispersion/air quality model, AERMOD.	Emissions data from the study *Characterizing Ammonia Emissions from Swine Farms in Eastern North Carolina: Part 1—Conventional Lagoon and Spray Technology for Waste Treatment (2008)* from the Barham and Moores farms. Dry deposition of NH_3 was modeled, not measured.	NH_3 flux measurements were made over waste lagoons using a dynamic flow-through chamber system.	This study has limited applicability to EEM development.

Table 3-4. Review of Dairy Lagoon Articles Received in Response to EPA's CFI

Year	Title - Author(s)	Brief Description	Relevant Data	Methodology	Possible Application for NAEMS
2006	*Quality Assured Measurements of Animal Building Emissions: Gas Concentrations* (Heber, Ni, Lim, Tao, Schmidt, Koziel, Beasley, Hoff, Nicolai, Jacobson, Zhang)	This study focuses on the methodology of measuring gas concentration and the difficulty in achieving the desired results for livestock houses due to their unique traits.	None	Not applicable	None. The article provides an overview of the emissions monitoring system and instrumentation but does not contain emissions or process data.
2007	*Modeling Atmospheric Transport and Fate of Ammonia in North Carolina—Part I: Evaluation of Meteorological and Chemical Predictions* (Wu, Krishnan, Zhang, Aneja)	This study discusses the application of EPA's Community Multiscale Air Quality (CMAQ) modeling system to study the deposition and fate of NH_3 emissions from activities. Part I of the study describes the model configurations, evaluation protocols, databases and the operational evaluation for meteorological and chemical predictions.	None	Not applicable	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
2007	*Modeling Atmospheric Transport and Fate of Ammonia in North Carolina—Part II: Effect of Ammonia Emissions on Fine Particulate Matter Formation* (Wu, Hu, Zhang, Aneja)	This study discusses the application of EPA's CMAQ model to study the deposition and fate of NH_3 emissions from agricultural activities. Part II of the study describes the sensitivity simulations applied to various emission scenarios.	None	Not applicable	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
2008	*Ammonia Assessment from Agriculture: U.S. Status and Needs* (Aneja, Blunden, James, Schlesinger, Knighton, Gilliam, Jennings, Niyogi, Cole)	This article summarizes recent research on agricultural air quality and describes best management practices for reducing NH_3 emissions.	None	Not applicable	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
2008	*Auditing and Assessing Air Quality in Concentrated Feeding Operations* (Cole, Todd, Auvermann, Parker)	This paper discusses AFO emissions and the current air quality regulations and techniques for measuring and quantifying emissions.	None	Not applicable	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.

Table 3-4. Review of Dairy Lagoon Articles Received in Response to EPA's CFI

Year	Title - Author(s)	Brief Description	Relevant Data	Methodology	Possible Application for NAEMS
2008	*Commentary: Farming Pollution* (Aneja, Schlesinger, Erisman)	This article provides commentary on the U.S. efforts to regulate farms. It provides general information related to agricultural emissions and the state of knowledge of processes and a comparison of U.S. regulations to European regulations.	None	Not applicable.	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
2009	*Ammonia Emissions and Animal Agriculture* (Gay, Knowlton)	This article provides general information regarding AFO emissions and the effects of farming on pollution.	None	Not applicable	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
2009	*Does Animal Feeding Operation Pollution Hurt Public Health? A National Longitudinal Study of Health Externalities Identified By Geographic Shifts In Livestock Production* (Sneeringer)	This article discusses an epidemiological study that assessed the relationship between livestock farming and infant mortality.	None	Not applicable	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
2009	*Effects of Agriculture upon the Air Quality and Climate: Research, Policy, and Regulations* (Aneja, Schlesinger, Erisman)	This article describes the state of the science and how research can be improved.	None	Not applicable	None. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
2009	*2008 Dairy Emissions Study - Summary of Dairy Emission Factors and Emission Estimation Procedures* (Schmidt, Card)	This study measured air emissions from the surface of feed (total mixed ration), silage, and the corral at two dairies in CA. Emissions for the entire dairy were calculated based on surface areas, and diurnal/seasonal variations in emission rates for each surface.	None	This study used an EPA surface emissions isolation flux chamber to measure emissions from the various surfaces at the dairies. TNMNEO emissions were quantified using South Coast Air Quality Management District (SCAQMD) Method 25.3. NH_3 was measured using SCAQMD Method 207.1.	None. This study did not measure emissions from a lagoon.

3.2.2 Previous Literature Searches

Beginning in 2001, the EPA conducted several literature searches using the Agricultural Online Access (AGRICOLA) bibliographic database to identify data and information that were relevant to support a preliminary investigation into air pollution from large AFOs (see the EPA's *Emissions From Animal Feeding Operations* (draft, August 15, 2001)). The EPA also conducted literature searches to support development of the EPA's National Emissions Inventory (NEI) for NH_3 emissions from animal agricultural operations.

Table 3-5 and Table 3-6 lists additional articles and publications pertaining to swine and dairy operations, respectively, that the EPA identified through literature searches it conducted prior to the CFI. Articles that were common to both the CFI and previous literature searches are reported in Table 3-3 and Table 3-4 only. As shown in the Table 3-5 and Table 3-6, none of the articles previously obtained by the EPA to support emissions factor development used remote sensing techniques to measure lagoon emissions. Consequently, none of the articles were applicable for EEM development.

Table 3-5. Review of Swine Lagoon Articles Obtained by Previous EPA Literature Searches

Date	Title	Author	Possible Application for NAEMS
1971	*Desorption of Ammonia from Anaerobic Lagoons*	Koelliker, Miner	None. This study used a mass balance method to estimate the amount of NH_3 lost from the storage lagoon due to desorption. Emissions were estimated, rather than measured. Consequently, these data were not considered for use in EEM development.
1981	*Storage of Piggery Slurry*	Williams, Evans	None. This study examined the changes in composition of swine manure over time. Changes in organic matter solubility, NPK value, BOD, odor offensiveness, and malodorous compounds were examined. Emissions rates were not examined. Consequently, these data were not considered for use in EEM development.
1988	*Odors of Swine Waste Lagoons*	Hammond, Heppner, Smith	None. This study focused on the effect of lagoon temperature on odor concentrations of emissions. Consequently, these data were not considered for use in EEM development.
1993	*Factors Affecting Ammonium Concentrations in Slurry from Fattening Pigs*	Aarnink, A., Hoeksma, and Ouwerkerk	None. This study investigated the effect of various factors (animal growth, feed protein content, feed intake, water-feed ratio) on the NH_3 content in the MESPRO model (NH_3 content of the swine waste slurry). This study did not measure NH_3 emissions. Consequently, these data were not considered for use in EEM development.
1994	*Biotreatment of Swine Manure by intensive Lagooning during winter*	Noue, Sevrin-Reyssac, Mariojouls, Marcel, and Sylvestre	None. This study investigated removal of nitrogen and phosphorus compounds from swine manure in a lagoon system consisting of algae ponds daphnid ponds and a polishing fish pond. This experiment did not evaluate emissions from the lagoon surface. Consequently, these data were not considered for use in EEM development.
1997	*Effect of Lagoon Aeration on Odor Emissions from a Swine Grow-Finish Facility*	Heber	None. This study focused on the methodology for collecting and analyzing samples for odor concentrations. Consequently, these data were not considered for use in EEM development.

Table 3-5. Review of Swine Lagoon Articles Obtained by Previous EPA Literature Searches

Date	Title	Author	Possible Application for NAEMS
1998	*Ammonia Emissions from Swine Waste Lagoons in the Southeastern U.S. Coastal Plains*	Harper and Sharpe	None. This study investigated the effect of diurnal variations, seasonal variations, lagoon ammonium (NH_4^+) concentration, lagoon acidity, temperature, and wind turbulence on NH_3 emissions from two lagoons on two farms using gas sensors and meteorological instruments placed over the lagoons. NH_3 concentrations were determined using air drawn through gas washing bottles; wash water was examined using colorimetric analysis.
1999	Ammonia Emissions from Swine Waste Operations in North Carolina	Aneja, Bunton, Chauhan, Malik, Walker, Li	None. This study examined NH_3 emissions from six swine lagoons using a dynamic chamber system, with the intent of parameterizing the NH_3-N flux with respect to lagoon physicochemical properties for use in air quality modeling. Water temperature and water NH_3-N content were monitored, as were NH_3-N emissions. NH_3 emission rate was determined as a factor of temperature and water NH_3-N content. The article did not specify the method by which NH_3 in the flux chamber was measured.
1999	*Odor Emissions from a Swine Finishing Facility with a Surface-Aerated Lagoon*	Heber and Ni	None. This study focused on the methodology for collecting and analyzing samples for odor concentrations. Consequently, these data were not considered for use in EEM development.
1999	*Slurry Covers to Reduce Ammonia Emissions and Odour Nuisance*	Hornig, Turk, Wanka	None. This study investigated the effect of floating covers on swine waste lagoons at four farms as a means of reducing NH_3 and odor emissions. This study did not quantify NH_3 emissions. Consequently, these data were not considered for use in EEM development.
1999	*Stakeholders Feedlot Air Emission Data Collection Project: Final Report, 12/30/99*	Jacobson	None. This report presents the results of monitoring for NH_3, H_2S, and odor at five swine farms in Minnesota. Emissions were estimated from short-term grab samples. NH_3 concentrations were quantified using colorimetric tubes. H_2S concentrations were measured using a Jerome gas analyzer.

Table 3-5. Review of Swine Lagoon Articles Obtained by Previous EPA Literature Searches

Date	Title	Author	Possible Application for NAEMS
2000	*Characterization of Atmospheric Ammonia Emissions from Swine Waste Storage and Treatment Lagoons*	Aneja, Chauhan, Walker	None. This study investigated the seasonal variation in NH_3 flux from a swine waste storage lagoon using a dynamic chamber system. Lagoon physicochemical properties (water temperature, NH_3-N content, and pH) were monitored. NH_3 flux was measured with a Measurement Technologies 1000N NH_3-to-NO converter and API Model 200 chemiluminescence based NO monitor.
2000	*A Comparison of the Performance of Three Swine Waste Stabilization Systems*	Martin, J.	None. This study compared emissions of CH_4, CO_2, and NH_3 from the surface of three lagoons (covered anaerobic lagoon with CH_4 capture, open lagoon with ozone injection, and open lagoon. The method for determining emissions was not discussed. Consequently, these data were not considered for use in EEM development.
2000	*Hydrogen Sulfide from Lagoons and Barns*	Secrest, C.	None. This one-page summary shows H_2S emissions from seven swine lagoons. The emissions were calculated based on results from four different estimation methods. Two of the methods were based on mass balance and two were based on measurements made in independent studies. Consequently, these data were not considered for use in EEM development.
2001	*Measurement and Analysis of Atmospheric Ammonia Emissions from Anaerobic lagoons*	Aneja, Bunton, Walker, Malik	None. This study examined the NH_3 flux from six anaerobic swine lagoons using a dynamic flux chamber system. Lagoon physicochemical properties (water temperature, NH_3-N content, and pH) were monitored. NH_3 flux was measured with a Measurement Technologies 1000N NH_3 to NO converter and API Model 200 chemiluminescence based NO monitor. The study examined the NH_3 flux as a factor of daily diurnal time, pH and water N content.
2001	*Measurement and Modeling of Ammonia Emissions at Waste Treatment Lagoon-Atmospheric Interface*	Aneja, Malik, Tong, Kang, Overton	None. This study investigated two mass transport models as predictive models for predicting NH_3 emissions from swine waste lagoons The model results were compared with NH_3 emission experiments at a swine lagoon. NH_3 emissions were measured using flux chambers.

Table 3-5. Review of Swine Lagoon Articles Obtained by Previous EPA Literature Searches

Date	Title	Author	Possible Application for NAEMS
2002	*Methane Emissions from Swine lagoons in Southeastern U.S.*	Sharpe, Harper, Byers	None. This study measured the emissions of CH_4 from two swine waste lagoons. Meteorological conditions were monitored. Methane concentrations were measured above the lagoon at two heights with tunable diode laser spectroscopy. Lagoon water was sampled for total organic carbon, volatile solids, and pH.
2003	*Atmospheric Pollutants and Trace Gases--Odor and Gas Release from Anaerobic Treatment Lagoons for Swine Manure*	Lim, Heber, Ni, Sutton, Shao	None. This study analyzes the effect of lagoon loading rate on emissions of NH_3, H_2S, CO_2, SO_2 and NO from a swine lagoon using a buoyant conductive flux chamber for sampling concentrations above the lagoon surface, downwind air samplers for sampling concentrations on the lagoon berm, and a gas chromatograph for analyzing gas samples.

Table 3-6. Review of Dairy Lagoon Articles Obtained by Previous EPA Literature Searches

Date	Title	Author	Possible Application for NAEMS
1991	*Nitrogen Concentration Variability in Dairy-Cattle Slurry Stored in Farm Tanks*	Patni, Jui	None. This study examined the nitrogen content of slurry stored in tanks over a period of several months. Emissions were not measured. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
1995	*Separation of Manure Solids from Simulated Flushed Manures by Screening or Sedimentation.*	Powers, Montoya, Van Horn, Nordstedt, Bucklin	None. This study investigated the amount of solids, nitrogen, phosphorus that could be removed from cow feces and urine diluted in water using screening versus various precipitation methods. NH_3 emissions from lagoons were not measured. The article does not contain emissions or process data that could be used to supplement the NAEMS data.
1999	*Wheat Straw Cover for Reducing Ammonia and Hydrogen Sulfide Emissions from Dairy Manure Storage*	Xue, S.K., S. Chen, R.E. Hermanson	None. This study investigated the use of wheat straw cover for reducing NH_3 and H_2S emissions from liquid manure. The study was conducted in the laboratory and using pilot-scale tanks. The method used to analyze the gas samples was not specified in the article.
2002	*Particulate and Dissolved Phosphorus Chemical Separation and Phosphorus Release from Treated Dairy Manure*	Dao, Daniel	None. This study investigated the effect of water treatment polymers and phosphorus immobilizing chemicals on particulate and dissolved reactive phosphorus reduction mechanisms in dairy manure. The article does not contain emissions or process data that could be used to supplement the NAEMS data.

4.0 NAEMS DATA PREPARATION

This section provides an overview of the data assessment procedures followed by the NAEMS researchers in collecting the emissions and process parameter data from the swine and dairy lagoon monitoring sites and the procedures followed by the EPA in preparing the data for use in development of EEMs for lagoons.

Section 4.1 discusses the Quality Assurance/Quality Control (QA/QC) procedures outlined in the NAEMS QAPP to ensure collection of high-quality emissions and process data. Section 4.2 summarizes the steps the EPA followed to process and review the data submitted to the EPA prior to developing the lagoon EEMs. Section 4.3 compares the design and operating parameters and reported emissions of each site.

4.1 NAEMS Data Assessments

4.1.1 QA/QC Procedures

The NAEMS researchers followed QA/QC procedures throughout the data collection and preliminary data analyses processes. The NAEMS researchers developed QAPPs, SOPs for sampling systems and monitoring instruments and site-specific monitoring plans (SMPs) and provided training for on-site operators and producers. Appendix A contains the QAPPs, Appendix B contains the SMPs for each monitoring site and Appendix C contains the SOPs.

Remote sensing instruments, synthetic open-path sampling systems, and gas analyzers underwent initial and periodic calibration, bias and precision checks and were corrected if they failed the QC checks. The frequency of each check/calibration event was dependent on the type of instrument and on the site investigators. For example, the calibration checks were conducted daily, at the beginning and end of each measurement period, every 180 days at Purdue Air Monitoring Lab (PAML), annually or a combination of these intervals.

All of the monitoring sites were equipped with data acquisition (DAQ) systems that allowed on-site operators, and other authorized personnel via high-speed Internet connection, to view the measured data and parameter values daily through real-time computer displays. The DAQ systems also generated email notifications for project personnel when monitored parameter values were outside of preset ranges.

The NAEMS researchers also used control charts extensively in QA/QC procedures to assess data quality and measurement variability and to evaluate long-term trends in the instrument/equipment performance. The control charts provided a graphical means of determining whether the measured parameters were within acceptable upper and lower control limits. Data values outside the control limits triggered corrective actions by site operators to maintain data quality. The control charts were generated on site using Microsoft ® Excel templates to provide a real-time assessment of the data quality.

Measurement data recorded at each site were transferred to PAML each day for review and evaluation. The NAEMS field operations staff used custom-designed software to apply flags to measurement data that were considered invalid or outliers and to calculate emissions rates for the lagoons.

4.1.2 Data Validation

In general, the NAEMS researchers invalidated measurement data if the data values were:

- Unreasonably low or high when compared to normal ranges if there was supporting evidence that the data value is not correct (e.g., lagoon temperature sensor producing a reading of less than -40 ° C).
- Obtained during system installation, testing or maintenance during which uncorrectable errors might be introduced.
- Obtained when a sensor or instrument was proven to be malfunctioning (e.g., unstable).
- Obtained during calibration or precision check of a sensor or instrument and before the sensor or instrument reached equilibrium after the check.
- Obtained when the data acquisition and control hardware and/or software were not functioning correctly.

Data that the NAEMS researchers deemed invalid were retained in the preprocessed data sets. However, the EPA did not use the flagged data to calculate pollutant emissions rates.

For averaged data, data were invalidated to avoid errors introduced into calculated mean values due to partial-data days (e.g., only a few hours of valid data) that would result in biased time weights:

- Daily means were invalidated if less than 75 percent of the 30-minute intervals recorded during that day were valid.
- Quarters, or monitoring periods, achieved the 75-percent completeness criteria if 7.5 days of valid daily measurements were obtained during a 10-day period.
- Three of the four quarters each year must attain the quarterly completeness criteria for the year to be adequately represented in the study results.

4.1.3 Data Completeness

As specified in the open-source QAPP, the goal of the NAEMS was to measure emissions from lagoons at nine farms that represented the dominant manure management systems for each animal type. At each of the nine sites, data were collected on a continuous and intermittent basis.

4.1.3.1 Continuous Data

For the swine and dairy sites, the continuous sampling program under the NAEMS consisted of long- and short-term monitoring. Under long-term monitoring, continuous measurements were taken over a one-year period at a single site for each animal type (site IN4A for swine and site IN5A for dairy). The short-term monitoring consisted of continuous measurements taken at multiple sites on a rotating sampling schedule over a two-year period. Meteorological and lagoon liquid data were recorded as 5-minute average values, and emissions data were reported as 30-minute averages. These data were aggregated by the NAEMS researchers to obtain daily values.

Consistent with the EPA's *Guidance for Quality Assurance Project Plans* (EPA QA/G-5), data completeness is defined as the measure of the amount of valid data obtained from a measurement system, compared to the amount of data that was expected to be obtained under normal conditions. Data completeness is expressed as the percent of valid data obtained from the measurement system. For data to be considered valid, they must meet all the acceptance criteria. As specified in the NAEMS QAPP, a valid day for a pollutant or process parameter was one in which more than 75 percent of the 30-minute measurement values recorded were valid (i.e., 36 half hours in the day passed all QA checks). The long-term monitoring sites had a second completeness goal to obtain valid emissions measurements for 75 percent of the total number of monitoring days on site (273 days would meet daily completeness criteria). For the short-term monitoring sites, the second completeness goal was to obtain emissions measurements for 10 days per quarter. The duration of the sampling events at any one site was dictated by weather conditions with a maximum duration of 21 days per quarter. The maximum duration was intended to ensure that the data quality indicator of 75 percent completeness of 10 days per quarter would not be prevented by adverse weather. Table 4-1 summarizes the monitoring schedule, the number of valid emissions days and data completeness by site. Table 4-1 also shows the quarter and season corresponding to each monitoring day. To simplify the seasonal assignments, if the monitoring period overlapped with another season by less than 15 days, the EPA assigned the monitoring period to the season that had the majority of days. The EPA assigned monitoring periods that overlapped the next season by more than 15 days as mixed-season periods (e.g., Fall/Winter). In the development of EEMs, the EPA considered all valid data days regardless of whether or not the NAEMS completeness goal was achieved.

At all of the short-term monitoring sites, the measurement instrumentation was on site for at least once each quarter/season over the 2-year monitoring period. For NH_3 emissions calculated using the RPM model, the 75-percent completeness criteria was achieved only at site OK4A. The final reports submitted to EPA do not discuss data completeness and do not provide detailed reasons for why the completeness goals were not achieved.

Table 4-1. Reported Number of Valid Emissions Days by Site

Site ID	Monitoring Period	Start Date	End Date	Quarter	Season[a]	Days On Site	No. of Valid Emissions Days				Data Completeness			
							NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratio	NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratio
IN4A[b]	1	7/1/2007	8/31/2007	3	Summer	61	0	7	0	0	0%	11%	0%	0%
	2	9/1/2007	11/30/2007	4	Fall	90	0	19	0	0	0%	21%	0%	0%
	3	12/1/2007	2/28/2008	1	Winter	89	0	0	6	0	0%	0%	7%	0%
	4	3/1/2008	5/31/2008	2	Spring	91	0	9	3	0	0%	10%	3%	0%
	5	6/1/2008	7/14/2008	3	Summer	43	1	2	6	0	2%	5%	14%	0%
NC4A	2	10/4/2007	10/22/2007	4	Fall	18	0	0	0	0	0%	0%	0%	0%
	3	1/29/2008	2/11/2008	1	Winter	13	0	2	0	0	0%	20%	0%	0%
	4	3/31/2008	4/16/2008	2	Spring	16	0	6	3	0	0%	60%	30%	0%
	5	8/13/2008	9/2/2008	3	Summer	20	0	1	3	0	0%	10%	30%	0%
	6	9/4/2008	9/23/2008	4	Fall	19	1	6	5	2	10%	60%	50%	20%
	7	1/14/2009	2/2/2009	1	Winter	19	0	1	1	0	0%	10%	10%	0%
	8	4/28/2009	5/11/2009	2	Spring	13	0	0	0	6	0%	0%	0%	60%
	9	7/1/2009	7/21/2009	3	Summer	20	3	0	0	5	30%	0%	0%	50%
OK4A	1	6/27/2007	8/29/2007	3	Summer	63	9	12	0	0	**90%**	**120%**	0%	0%
	2	11/7/2007	11/27/2007	4	Fall	20	0	2	0	0	0%	20%	0%	0%
	3	11/28/2007	12/18/2007	1	Winter	20	0	5	0	0	0%	50%	0%	0%
	4	4/23/2008	5/6/2008	2	Spring	13	0	2	2	0	0%	20%	20%	0%
	6	10/1/2008	10/15/2008	4	Fall	14	2	7	7	1	20%	70%	70%	10%
	7	1/8/2009	1/27/2009	1	Winter	19	0	0	2	0	0%	0%	20%	0%
	8	4/1/2009	4/21/2009	2	Spring	20	10	10	8	0	**100%**	**100%**	**80%**	0%
	9	6/25/2009	7/14/2009	3	Summer	19	9	8	5	1	**90%**	**80%**	50%	10%
IA3A	2	8/30/2007	9/26/2007	4	Fall	27	0	2	0	0	0%	20%	0%	0%
	3	12/19/2007	1/15/2008	1	Winter	27	0	7	0	0	0%	70%	0%	0%
	4	5/14/2008	6/4/2008	2	Spring	21	0	3	0	0	0%	30%	0%	0%
	5	6/4/2008	6/25/2008	2	Spring	21	0	5	1	0	0%	50%	10%	0%
	6	11/13/2008	11/25/2008	4	Fall	12	0	6	3	0	0%	60%	30%	0%
	7	11/25/2008	12/16/2008	4/1	Fall/ Winter	21	0	4	1	0	0%	40%	10%	0%
	8	4/8/2009	4/23/2009	2	Spring	15	0	3	1	0	0%	30%	10%	0%
	9	7/28/2009	8/17/2009	3	Summer	20	4	2	0	1	40%	20%	0%	10%
NC3A	2	10/24/2007	11/7/2007	4	Fall	14	0	0	0	0	0%	0%	0%	0%
	3	2/13/2008	3/5/2008	1	Winter	21	0	2	1	0	0%	20%	10%	0%
	4	3/6/2008	3/26/2008	2	Spring	20	0	1	1	0	0%	10%	10%	0%
	6	9/25/2008	10/14/2008	4	Fall	19	0	0	0	0	0%	0%	0%	0%

Table 4-1. Reported Number of Valid Emissions Days by Site

Site ID	Monitoring Period	Start Date	End Date	Quarter	Season[a]	Days On Site	No. of Valid Emissions Days				Data Completeness			
							NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratio	NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratio
	7	2/4/2009	2/23/2009	1	Winter	19	2	1	0	1	20%	10%	0%	10%
	8	5/12/2009	6/2/2009	2	Spring	21	0	0	0	0	0%	0%	0%	0%
	9	6/2/2009	6/22/2009	3	Summer	20	1	0	0	1	10%	0%	0%	10%
	10	9/24/2009	12/1/2009	4	Fall	68	0	4	4	0	0%	40%	40%	0%
OK3A	2	8/30/2007	9/18/2007	4	Fall	19	4	8	0	0	40%	**80%**	0%	0%
	3	1/24/2008	2/19/2008	1	Winter	26	0	8	2	0	0%	**80%**	20%	0%
	4	5/7/2008	5/29/2008	2	Spring	22	1	6	7	0	10%	**60%**	70%	0%
	5	5/29/2008	6/10/2008	3	Summer	12	0	8	2	0	0%	**80%**	20%	0%
	6	11/5/2008	12/2/2008	4	Fall	27	0	7	5	0	0%	**70%**	50%	0%
	7	12/2/2008	12/16/2008	1	Winter	14	0	3	0	0	0%	30%	0%	0%
	8	4/23/2009	5/14/2009	2	Spring	21	0	4	5	0	0%	40%	50%	0%
	9	7/15/2009	8/4/2009	3	Summer	20	4	6	3	2	40%	60%	30%	20%
IN5A[b]	6	9/11/2008	11/30/2008	4	Fall	80	8	22	5	0	10%	28%	6%	0%
	7	12/1/2008	2/28/2009	1	Winter	89	0	15	4	0	0%	17%	4%	0%
	8	3/1/2009	5/31/2009	2	Spring	91	8	27	10	1	9%	30%	11%	1%
	9	6/1/2009	8/17/2009	3	Summer	77	2	9	5	0	3%	12%	6%	0%
WA5A	3	2/25/2008	3/12/2008	2	Spring	13	0	0	0	0	0%	0%	0%	0%
	4	3/12/2008	3/26/2008	2	Spring	14	0	0	0	0	0%	0%	0%	0%
	5	8/8/2008	9/3/2008	3	Summer	26	0	0	0	0	0%	0%	0%	0%
	6	9/3/2008	9/26/2008	4	Fall	23	0	0	0	0	0%	0%	0%	0%
	8	5/18/2009	6/4/2009	2	Spring	16	0	1	0	0	0%	10%	0%	0%
	9	6/4/2009	6/20/2009	3	Summer	15	0	0	0	0	0%	0%	0%	0%
WI5A	1	7/18/2007	8/28/2007	3	Summer	41	0	0	0	0	0%	0%	0%	0%
	2	11/13/2007	11/28/2007	4	Fall	15	0	4	0	0	0%	40%	0%	0%
	3	11/28/2007	12/18/2007	1	Winter	20	0	2	0	0	0%	20%	0%	0%
	4	4/23/2008	5/13/2008	2	Spring	20	0	2	0	0	0%	20%	0%	0%
	5	6/25/2008	7/14/2008	3	Summer	19	0	3	1	0	0%	30%	10%	0%
	6	10/21/2008	11/11/2008	4	Fall	21	0	5	2	0	0%	50%	20%	0%
	7	12/17/2008	1/7/2009	1	Winter	21	0	1	0	0	0%	10%	0%	0%
	8	3/10/2009	4/7/2009	2	Spring	28	0	5	3	0	0%	50%	30%	0%

[a] The seasons are defined in the NAEMS open-source QAPP as follows: Fall= September, October, November; Winter= December, January, February; Spring= March, April, May; Summer= June, July, August.
[b] Long-term monitoring was conducted at this site.

4.1.3.2 Intermittent Data

The intermittent data collected under the NAEMS consisted of lagoon composition data, lagoon coverage and animal inventory values. For the lagoon composition data, the QAPP states that samples of manure were to be collected at the lagoons every measurement period (i.e., four times per year). Table 4-2 summarizes the manure sampling schedule for each site, the availability of composition data and whether or not the sampling data coincided with the monitoring periods for each site. Although most of the sites reported liquid composition data once per measurement period, the sampling dates rarely coincided with a NAEMS monitoring date. Considering alignment with the NAEMS monitoring dates, the data completeness goal for liquid composition data of quarterly sampling was not achieved at any of the NAEMS monitoring sites. Liquid composition data were not collected at site WA5A and the data for sites IN4A, IN5A and WI5A were reported as ranges and the sampling dates were not provided.

Completeness criteria for the lagoon coverage and animal inventory values were not specified in the QAPP. Table 4-3 shows the dates that observations were recorded for lagoon coverage. Table 4-4 shows the dates for which an animal inventory value were provided.

Table 4-2. Summary of Intermittent Data

Site	Sampling Date	Number of Samples	Content Data Available?					Sampling Date Coincident w/NAEMS Monitoring Date?
			pH	Nitrogen	Solids	NH_3	Sulfur	
IN4A	The final report contained ranges of composition data without sampling dates.							Unknown
NC4A	12/1/2006	1	Y	Y	N	N	Y	N
	1/24/2007	1	Y	Y	N	N	Y	N
	3/22/2007	1	Y	Y	N	N	Y	N
	4/19/2007	1	Y	Y	N	N	Y	N
	6/15/2007	1	Y	Y	N	N	Y	N
	8/1/2007	1	Y	Y	N	N	Y	N
	9/25/2007	1	Y	Y	N	N	Y	N
	11/20/2007	1	Y	Y	N	N	Y	N
	1/24/2008	1	Y	Y	N	N	Y	N
	5/16/2008	1	Y	Y	N	N	Y	N
	7/9/2008	1	Y	Y	N	N	Y	N
	8/26/2008	1	Y	Y	N	N	Y	N
	10/15/2008	1	Y	Y	N	N	Y	N
	2/18/2009	1	Y	Y	N	N	Y	N
	4/3/2009	1	Y	Y	N	N	Y	N
	4/3/2009	1	Y	Y	N	N	Y	N
	5/25/2009	1	Y	Y	N	N	Y	N
	5/26/2009	1	Y	Y	N	N	Y	N
	7/17/2009	1	Y	Y	N	N	Y	Y
	8/31/2009	1	Y	Y	N	N	Y	N
	12/16/2009	1	Y	Y	N	N	Y	N

Table 4-2. Summary of Intermittent Data

Site	Sampling Date	Number of Samples	Content Data Available?					Sampling Date Coincident w/NAEMS Monitoring Date?
			pH	Nitrogen	Solids	NH_3	Sulfur	
OK4A	11/1/2008	1	Y	Y	Y	Y	Y	N
IA3A	6/4/2008	3	Y	N	Y	Y	N	Y
	4/9/2009	3	Y	Y	Y	Y	N	Y
	4/15/2009	3	Y	Y	Y	Y	N	N
	7/29/2009	3	Y	Y	Y	Y	Y	Y
NC3A	12/1/2006	1	Y	Y	N	N	Y	N
	1/24/2007	1	Y	Y	N	N	Y	N
	3/22/2007	1	Y	Y	N	N	Y	N
	4/19/2007	1	Y	Y	N	N	Y	N
	6/15/2007	1	Y	Y	N	N	Y	N
	8/6/2007	1	Y	Y	N	N	Y	N
	9/27/2007	1	Y	Y	N	N	Y	N
	11/20/2007	1	Y	Y	N	N	Y	N
	1/23/2008	1	Y	Y	N	N	Y	N
	3/25/2008	1	Y	Y	N	N	Y	Y
	5/16/2008	1	Y	Y	N	N	Y	N
	7/10/2008	1	Y	Y	N	N	Y	N
	8/26/2008	1	Y	Y	N	N	Y	N
	10/15/2008	1	Y	Y	N	N	Y	N
	12/16/2008	1	Y	Y	N	N	Y	N
	2/18/2009	1	Y	Y	N	N	Y	Y
	3/31/2009	1	Y	Y	N	N	Y	N
	5/21/2009	1	Y	Y	N	N	Y	Y
	7/13/2009	1	Y	Y	N	N	Y	N
	9/1/2009	1	Y	Y	N	N	Y	N
OK3A	11/28/2007	1	Y	Y	Y	Y	Y	N
	11/19/2008	1	Y	Y	Y	Y	Y	Y
	7/17/2009	1	Y	Y	Y	Y	Y	Y
IN5A	The final report contained ranges of composition data without sampling dates.							Unknown
WA5A	None	N	N	N	N	N	N	Not applicable
WI5A	The final report contained ranges of composition data without sampling dates.							Unknown

Table 4-3. Dates of Lagoon Coverage Observations

Site	Observation Date
IN4A	07/20/2007 - 03/12/2008, 07/24/2007 - 03/17/2008, 08/08/2007 - 03/20/2008, 08/10/2007 - 03/21/2008, 08/15/2007 - 04/07/2008, 08/16/2007 - 04/08/2008, 08/17/2007 - 04/14/2008, 08/20/2007 - 04/15/2008, 08/22/2007 - 04/30/2008, 08/23/2007 - 05/01/2008, 08/30/2007 - 05/12/2008, 09/04/2007 - 05/20/2008, 09/05/2007 - 05/21/2008, 09/12/2007 - 05/27/2008, 09/19/2007 - 05/28/2008, 09/20/2007 - 06/05/2008, 10/04/2007 - 06/19/2008, 10/18/2007 - 06/20/2008, 10/23/2007 - 06/24/2008, 11/01/2007 - 07/01/2008, 11/13/2007 - 07/02/2008, 11/15/2007 - 07/03/2008, 01/29/2008 - 07/08/2008, 01/30/2008 - 07/14/2008, 01/31/2008 - 07/15/2008, 02/12/2008 - 07/16/2008, 3/5/2008

Table 4-3. Dates of Lagoon Coverage Observations

Site	Observation Date
NC4A	10/2/2007, 10/3/2007, 10/4/2007, 10/5/2007, 10/22/2007, 1/29/2008, 1/30/2008, 1/31/2008, 2/1/2008, 2/11/2008, 2/12/2008, 3/31/2008, 4/1/2008, 4/2/2008, 4/14/2008, 4/15/2008, 4/16/2008, 4/17/2008, 8/12/2008, 8/13/2008, 8/14/2008, 8/26/2008, 9/2/2008, 9/3/2008, 9/4/2008, 9/22/2008, 9/23/2008, 1/14/2009, 1/15/2009, 1/16/2009, 2/2/2009, 2/3/2009, 4/28/2009, 4/29/2009, 5/11/2009, 5/12/2009, 6/30/2009, 7/1/2009, 7/21/2009, 7/22/2009
OK4A	6/27/2007, 6/28/2007, 7/9/2007, 7/18/2007, 7/19/2007, 7/30/2007, 7/31/2007, 8/1/2007, 8/2/2007, 8/3/2007, 8/4/2007, 8/18/2007, 8/16/2007, 8/28/2007, 8/29/2007, 11/7/2007, 11/8/2017, 11/27/2007, 11/28/2007, 12/18/2007, 12/19/2007, 4/23/2008, 4/24/2008, 5/6/2008, 10/1/2008, 10/2/2008, 10/14/2008, 10/15/2008, 1/7/2009, 1/8/2009, 1/27/2009, 4/1/2009, 4/2/2009, 4/21/2009, 6/23/2009, 6/24/2009, 6/25/2009, 7/14/2009
IA3A	08/30/2007 - 06/24/2008, 08/31/2007 - 11/12/2008, 09/17/2007 - 11/13/2008, 09/18/2007 - 11/14/2008, 09/26/2007 - 11/25/2008, 12/19/2007 - 12/15/2008, 12/20/2007 - 12/16/2008, 01/14/2008 - 04/08/2009, 01/15/2008 - 04/09/2009, 05/14/2008 - 04/22/2009, 05/15/2008 - 04/23/2009, 06/04/2008 - 07/27/2009, 06/05/2008 - 07/28/2009, 06/12/2008 - 07/29/2008, 06/23/2008 - 08/18/2009
NC3A	10/24/2007, 10/25/2007, 10/26/2007, 11/7/2007, 2/13/2008, 2/14/2008, 3/6/2008, 3/26/2008, 3/27/2008, 9/24/2008, 9/25/2008, 2/4/2009, 2/5/2009, 2/6/2009, 2/24/2009, 5/12/2009, 5/14/2009, 6/3/2009, 6/22/2009, 6/23/2009, 9/22/2009, 9/23/2009, 9/24/2009, 10/13/2009, 10/26/2009, 10/27/2009, 11/10/2009, 11/11/2009, 12/2/2009
OK3A	8/30/2007, 8/31/2007, 9/18/2007, 9/19/2007, 1/24/2008, 1/25/2008, 2/5/2008, 2/6/2008, 2/19/2008, 5/6/2008, 5/7/2008, 5/8/2008, 5/19/2008, 5/28/2008, 5/29/2008, 6/10/2008, 11/5/2008, 11/6/2008, 11/7/2008, 11/17/2008, 12/2/2008, 12/3/2008, 12/16/2008, 4/21/2009, 4/22/2009, 4/23/2009, 5/6/2009, 5/14/2009, 7/14/2009, 7/15/2009, 7/16/2009, 8/4/2009
IN5A	9/11/2008, 9/16/2008, 9/24/2008, 10/8/2008, 10/9/2008, 10/10/2008, 10/22/2008, 10/27/2008, 10/28/2008, 11/19/2008, 11/20/2008, 12/9/2008, 12/10/2008, 12/11/2008, 12/29/2008, 12/30/2008, 1/12/2009, 1/21/2009, 1/23/2009, 1/29/2009, 2/9/2009, 2/13/2009, 2/25/2009, 5/26/2009, 5/28/2009, 6/9/2009, 6/15/2009, 6/16/2009, 6/23/2009, 7/2/2009, 7/8/2009, 7/9/2009, 7/10/2009, 7/15/2009, 7/16/2009, 7/20/2009, 7/22/2009, 7/23/2009, 8/12/2009, 8/18/2009
WA5A	2/25/2008, 2/26/2008, 2/27/2008, 2/28/2008, 2/29/2008, 3/12/2008, 3/13/2008, 3/26/2008, 8/8/2008, 8/9/2008, 8/10/2008, 8/11/2008, 8/12/2008, 8/13/2008, 9/4/2008, 9/26/2008, 5/19/2009, 5/20/2009, 5/21/2009, 6/4/2009, 6/19/2009
WI5A	7/17/2007, 7/18/2007, 7/19/2007, 8/6/2007, 8/7/2007, 8/9/2007, 8/28/2007, 11/13/2007, 11/14/2007, 11/27/2007, 11/28/2007, 11/29/2007, 12/17/2007, 12/18/2007, 4/24/2008, 5/12/2008, 5/13/2008, 6/24/2008, 6/25/2008, 7/14/2008, 7/15/2008, 11/11/2008, 12/17/2008, 12/18/2008, 1/6/2009, 1/7/2009, 3/9/2009, 3/10/2009, 4/6/2009, 4/7/2009

Table 4-4. Dates or Animal Inventory Records

Site[a]	Monitoring Period	Dates for Which an Inventory Value Was Provided[b]
OK4A	2	11/6/2007 – 11/28/2007
	3	11/28/2007 - 12/19/2007
	4	4/22/2008 - 5/6/2008
	6	9/30/2008 - 10/15/2008
	7	1/7/2009 – 1/27/2009
	8	4/1/2009 - 4/21/2009
	9	6/22/2009 - 7/14/2009
IA3A	2	8/30/2007 - 9/26/2007
	3	12/19/2007 -1/15/2008
	4	5/14/2008 - 6/4/2008
	5	6/4/2008-6/25/2008
	6	11/13/2008-11/25/2008
	7	11/25/2008 - 12/16/2008
	8	4/8/2009-4/23/2009
	9	7/28/2009 - 8/17/2009
NC3A	2	10/24/2007 - 11/7/2007
	3	2/13/2008 - 3/5/2008
	4	3/6/2008 - 3/26/2008
	6	9/25/2008 - 10/14/2008
	7	2/4/2009 - 2/23/2009
	8	5/12/2009 - 6/2/2009
	9	6/2/2009 - 6/22/2009
	10	9/24/2009 - 12/2/2009
OK3A	3	1/24/2008 - 2/19/2008
	4	5/7/2008 - 5/29/2008
	5	5/29/2008 - 6/10/2008
	6	11/4/2008 -12/3/2008
	7	12/3/2008 - 12/16/2008
	8	4/21/2009 - 5/14/2009
	9	7/14/2009 - 8/4/2009
IN5A	6	9/11/2008 -11/30/2008
	7	12/1/2008 - 2/28/2009
	8	3/1/2009 - 5/31/2009
	9	6/1/2009 - 8/17/2009
WA5A	3	2/25/2008 - 3/12/2008
	4	3/12/2008 - 3/26/2008
	5	8/8/2008 - 9/3/2008
	6	9/3/2008 - 9/26/2008
	8	5/18/2009 - 6/4/2009
	9	6/4/2009 - 6/20/2009
WI5A	1	7/17/2007 - 8/28/2007
	2	11/12/2007 - 11/29/2007
	3	11/29/2007 - 12/18/2007
	4	4/22/2008 - 5/13/2008
	5	6/24/2008 - 7/15/2008
	6	10/20/2008 - 11/11/2008
	7	12/17/2008 -1/7/2009
	8	3/09/2009 - 4/7/2009

[a] Inventory data were not provided for sites IN4A and NC4A.
[b] A single inventory value was reported for the time period.

4.2 EPA Assessments

4.2.1 Data Processing

The NAEMS researchers provided the daily emissions, lagoon conditions and ambient meteorological data to the EPA in the form of final site reports (pdf format). The EPA also received Microsoft Excel® spreadsheets that contained emissions and lagoon condition data reported as 30-minute average values and ambient meteorological data reported as 5-minute average values.

Due to the very limited number of daily NH_3 emissions values reported, the EPA prepared a database of half-hour values by combining the 30-minute emissions data and 5-minute meteorological data provided by the NAEMS researchers. To obtain values for each meteorological variable for a given half-hour, the EPA used the mean of the 5-minute data values within that half hour, as long as at least one of the values was not missing. Because it is unlikely that the value of any of these variables would vary enough within the consecutive 5-minute periods that comprise a half hour to significantly change the mean value of the variable during the half-hour, applying a completeness criterion to this aggregation would unnecessarily reduce an already limited data set. In the data aggregation, the EPA did not include any 30- or 5-minute data that had been flagged by the NAEMS researchers as invalid.

4.2.2 Data QA

The EPA developed a comprehensive list of the emissions, farm and lagoon operating parameters, meteorological data and other information that were expected to be submitted to the EPA by the NAEMS researchers based on the EPA's review of the QAPPs, SOPs and SMPs. As the final reports and data spreadsheets were received from the NAEMS researchers, the EPA compared the information received to the comprehensive list to identify missing information. After determining whether the data submittals to the EPA were complete and identifying missing data elements, the EPA verified that the units of measurement for the emissions and supporting data were consistent between the final reports and spreadsheet data files. In addition, the EPA assessed whether the units of measurement and the magnitude of emissions were consistent across the monitoring sites. The EPA prepared and provided summaries of the missing data elements to the NAEMS researchers.

The EPA's review identified that the daily and 30-minute emissions values reported by the NAEMS researchers contained negative values. Table 4-5 shows the number of negative emissions by pollutant and measurement method.

Table 4-5. Negative Emissions Values by Pollutant and Measurement Method

Pollutant	Measurement Method	No. of Valid Emissions Values		No. of Negative Emissions Values		Percent Negative	
		Daily	30-min	Daily	30-min	Daily	30-min
NH_3	RPM	69	17,292	0	134	0%	1%
	bLS	285	6,416	18	1,108	6%	17%
H_2S	Ratiometric	21	6,325	1	876	5%	14%
	bLS	117	3,963	2	816	2%	21%

After discussion with the NAEMS Science Advisor, the EPA determined that the negative emission values occurred due to drift in the instrument readings between calibrations. The EPA did not include the negative values when developing EEMs for swine and dairy lagoons to avoid possible complications with EEM development (e.g., the EEM predicting negative emissions) the negative values were withheld from the data sets used for EEM development. The EPA's review also identified a few instances (less than 2 percent) of zero emission values (i.e., instances where the upwind/background and downwind concentrations were the same). However, because the zero values were not the result of instrument drift, the EPA included the zero emissions values in the data sets used in the development of the EEMs for lagoons.

4.2.3 Emissions Data Completeness Assessment

The EPA assessed the data completeness of the average daily emissions values to verify the completeness goals specified in the lagoon QAPP. The final reports did not include assessments of data completeness. Based upon its analysis, the EPA confirmed that the completeness goal for the two long-term monitoring sites was not achieved. Additionally, the completeness goal for the short-term monitoring sites was achieved only at site OK4A.

The EPA looked at the seasonal distribution of the data to determine if any pollutant was under-represented during a particular season (see Table 4-6 and Table 4-7). The NAEMS QAPP for lagoons defined the seasons as follows:

- Spring: March 1 through May 31.
- Summer: June 1 through August 31.
- Fall: September 1 through November 30.
- Winter: December 1 through February 28.

Table 4-6. Number of Valid Emissions Days Available in the Spring and Summer

Site ID	Spring				Summer			
	NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratiometric	NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratiometric
IN4A	0	9	2	0	1	9	9	0
NC4A	0	6	3	0	3	1	2	1
OK4A	10	12	10	6	18	20	5	5
IA3A	0	6	1	0	4	7	1	1
NC3A	0	1	1	0	1	0	0	1
OK3A	1	11	12	0	4	13	5	2
IN5A	8	27	10	1	2	9	5	0
WA5A	0	0	0	0	0	1	0	0
WI5A	0	7	3	0	0	3	1	0

Table 4-7. Number of Valid Emissions Days Available in the Fall and Winter

Site ID	Fall				Winter			
	NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratiometric	NH_3 RPM	NH_3 bLS	H_2S bLS	H_2S Ratiometric
IN4A	0	19	0	0	0	0	4	0
NC4A	1	6	6	1	0	3	1	0
OK4A	2	10	7	2	0	4	2	0
IA3A	0	8	3	0	0	11	1	0
NC3A	0	4	4	0	2	3	1	1
OK3A	4	15	5	0	0	11	2	0
IN5A	8	16	4	0	0	21	5	0
WA5A	0	0	0	0	0	0	0	0
WI5A	0	11	2	0	0	1	0	0

The seasonal distributions of valid NH_3 and H_2S emissions were weighted towards the spring and summer. At the other sites, the seasonal distributions of the emissions data were relatively consistent across each season. However, during each season, the amount of H_2S emissions data collected was lower than the NH_3 emissions data collected.

4.3 Comparison of Lagoon Monitoring Sites

Table 4-8 summarizes the configuration of each of the lagoon monitoring sites. The EPA developed comparative statistics and graphs of emissions data for each site to determine if there were any notable differences or data anomalies among the sites at the process, location or emissions level. Each of these comparisons is discussed in more detail in the following sections.

Table 4-8. NAEMS Data for Swine and Dairy Lagoon Confinement Operations

Site	Animal Sector	Confinement Description	Unit Measured	Manure Management System
IN4A	Swine	Sow	Anaerobic lagoon	Deep pit (emptied once every two weeks)
OK4A	Swine	Sow	Anaerobic lagoon	Pull plug pit w/pit recharge (emptied weekly from the two gestation units and every 2.5 weeks from the farrowing unit)
NC4A	Swine	Gestation, farrowing, and breeding	Anaerobic lagoon	Pull plug pit w/pit recharge (emptied once every week)
IA3A	Swine	Grow/finish	Storage lagoon	Deep pit (emptied ~ every 10 weeks)
NC3A	Swine	Grow/finish	Anaerobic lagoon	Pull plug pit w/pit recharge (emptied daily)
OK3A	Swine	Grow/finish	Anaerobic lagoon	Pull plug pit w/pit recharge (emptied 3 times a week)
IN5A	Dairy Lagoon	Milking Operation	Anaerobic lagoon	Scrape (daily)
WA5A	Dairy Lagoon	Milking Operation	Upper lagoon	Freestall flush (4 times/day)
WI5A	Dairy Lagoon	Milking Operation	1st and 2nd stages of the 3-stage lagoon system	Freestall, flush (3 times/day) w/solids separation

4.3.1 Process-Level Comparison

Table 4-9 summarizes the design and operating parameters for the lagoon sites. For the dairy sites, WA5A is the largest operation with 5,600-cow capacity, which is a large farm for the industry according to the 2007 USDA Census of Agriculture. Average animal weight across the dairies is comparable.

For the swine breeding/gestation farms, IN4A is the smallest operation with a capacity of 1,400 sows, while OK4A has the largest capacity at 2,784 sows. All three sites represent some of the larger breeding/gestation operations, compared to other farms in the industry. Average sow weight at the farms is comparable.

The swine grow/finish operations have more variation in size with both IA3A and OK3A having just over 3,000-head capacity and NC3A with a capacity of 8,000 head. The average animal weight for these site represents the average weight over the entire growing cycle, where the animal starts out as an approximately 14 pound piglet and final market weight of 250 pounds or more.

Table 4-9. Design and Operating Parameters of the NAEMS Swine and Dairy Lagoon Sites

Site	Farm Capacity (hd)	Animal Sector	Average Animal Weight (lb)	Confinement Description
IN4A	1,400 (sows)	Swine	475 (sows)/ 8 (piglets)	Sow
OK4A	2,784	Swine	490	Sow
NC4A	2,000 (sows)	Swine	433 (sows)	Gestation, farrowing, and breeding
IA3A	3,840	Swine	150	Grow/finish
NC3A	8,000	Swine	135	Grow/finish
OK3A	3,024	Swine	170	Grow/finish
IN5A	2,600	Dairy	1,500	Milking Operation
WA5A	5,600	Dairy	1,400	Milking Operation
WI5A	1,700	Dairy	1,400	Milking Operation

*Weight information not provided

4.3.2 Comparison of Local Meteorological Conditions

Table 4-10 summarizes the reported site-specific ambient and confinement conditions for each monitoring site. Ambient temperature varied greatly across the sites, with a minimum temperature at IN5A of -29.20 °C (-20.57 °F) and a maximum observed temperature of 41.37 °C (106.47 °F). NC4A had the highest average temperature at 18.07 °C (64.52 °F) and IN5A had the lowest average temperature at 4.47 °C (40.04 °F). The relative humidity observed during the study ranged from 5.05 to 100 percent. Site NC3A had the highest average humidity at 76.35 percent and WA5A had the lowest reported average humidity at 43.13 percent.

Lagoon liquid temperatures ranged from -0.51 °C (31.08 °F) at IN4A to 21.95 °C (95.01 °F) at IN5A. Lagoon temperature data were not available for IA3A and WA5A for any of the NAEMS monitoring periods. Lagoon temperature values were also not available when the surface of the lagoon liquid froze, which explains why there few lagoon temperatures reported below 0 °C.

Table 4-10. Site-Specific Ambient and Lagoon Conditions

Monitoring Site	Ambient				Lagoon	
	Temperature (°C)		Relative Humidity (%)		Temperature (°C)	
	Average	Range [Min, Max]	Average	Range [Min, Max]	Average	Range [Min, Max]
IN4A	11.33	[-20.32, 33.59]	75.31	[17.03, 100.00]	16.53	[-0.51, 34.11]
NC4A	18.07	[-10.68, 33.90]	74.34	[18.74, 100.00]	20.61	[2.34, 35.01]
OK4A	12.88	[-12.77, 40.97]	57.35	[6.00, 100.00]	19.81	[2.62, 34.59]
IA3A	9.72	[-21.58, 32.05]	75.55	[13.09, 100.00]	No data available	No data available
NC3A	15.10	[-7.30, 34.60]	76.35	[14.69, 100.00]	18.31	[5.59, 31.21]
OK3A	13.05	[-15.01, 41.37]	58.12	[5.05, 100.00]	16.19	[0.65, 30.47]
IN5A	4.47	[-29.20, 29.94]	72.63	[16.64, 99.22]	21.95	[10.32, 29.17]
WA5A	17.72	[-2.10, 38.58]	43.13	[8.04, 95.38]	No data available	No data available
WI5A	4.61	[-26.61, 29.10]	72.74	[17.41, 100.00]	18.30	[4.06, 33.89]

4.3.3 Emissions-Level Comparison

Table 4-11 summarizes the emissions values determined using the various measurement techniques for each of the lagoons monitored during the NAEMS. The data presented in the tables include all valid daily average values as reported in the final site reports.

Overall, dairies had the lowest average daily NH_3 emissions, and average daily H_2S emissions varied considerably within the animal types. Overall, OK4A had the highest NH_3 emission of all the sites monitored, for either measurement method. There is no obvious difference in this site from the other sites that would suggest a noticeable difference in NH_3 emissions (e.g., larger capacity or larger lagoon). For H_2S emissions, OK4A and IN5A have noticeably higher emission levels that the other sites. Again, there are no distinguishable characteristics that set these sites apart from the others, except that these two sites have the highest number of emissions values available for the bLS method.

Table 4-11. Average Daily Emissions for by Site

Pollutant	Parameter	Site								
		IN4A	NC4A	OK4A	IA3A	NC3A	OK3A	IN5A	WA5A	WI5A
NH_3 RPM (kg/d)	No. of values	1	4	30	4	3	9	18	0	0
	Average	58.4	70.8	177.2	47.1	58.2	102.8	23.4	-	-
	Range	[58.4, 58.4]	[53.4, 102.8]	[65.0, 317.9]	[38.3, 52.8]	[28.3, 117.8]	[53.6, 184.4]	[13.3, 34.6]	-	-
NH_3 bLS (kg/d)	No. of values	37	16	46	32	8	50	73	1	22
	Average	24.8	62.4	171.8	9.9	28.3	55.4	17.1	16.7	9.9
	Range	[0.1, 67.9]	[20.3, 136.2]	[4.8, 375.4]	[1.0, 59.6]	[4.4, 50.0]	[0.6, 149.8]	[0.7, 75.3]	[16.7, 16.7]	[1.5, 34.9]
H_2S bLS (kg/d)	No. of values	15	12	24	6	6	24	24	0	6
	Average	0.4	0.3	8.4	2.2	2.0	6.3	0.0	-	0.4
	Range	[0.0, 2.2]	[0.0, 1.0]	[0.1, 25.3]	[0.0, 12.3]	[0.0, 10.4]	[0.1, 25.4]	[0.0, 42.9]	-	[0.0, 1.9]
H_2S Radiometric (kg/d)	No. of values	0	2	13	1	2	2	1	0	0
	Average	-	0.2	14.3	9.2	4.5	0.2	12.5	-	-
	Range	-	[0.1, 0.3]	[2.4, 37.8]	[9.2, 9.2]	[9.0, 9.0]	[0.2, 0.2]	[12.5, 12.5]	-	-

5.0 OVERVIEW OF NH3 EEM

This section presents the statistical analysis and procedures followed by the EPA in developing EEMs for NH_3 emissions from lagoons at swine and dairy operations. The EEMs produces a point prediction of NH_3 emissions for a given half-hour period, given the values of predictor variables characterizing the farm and lagoon configuration and ambient meteorology. The EEMs can be used to obtain both a point prediction and a 95-percent prediction interval for the total daily and annual emissions for a lagoon.

The amount of time-varying data (e.g., nitrogen content, ammonia content, pH) available to characterize the lagoon liquid at each site was very limited and including these limited data as predictor variables would have significantly reduced the overall size of the dataset available for EEM development. To maximize the number of NH_3 emissions measurements used to develop the EEM, the EPA considered a suite of static, farm-based predictor variables as surrogates for the time-varying data. These farm-based predictor variables describe the total live mass of animals on the farm, the surface area of the lagoon, and other characteristics of the facility that might indirectly provide information about precursor loading and potential for conversion to NH_3 emissions. However, because this information is not time-varying, there are only nine data values (one for each site) for each of the farm-based predictor variables available for quantifying both the effects of these variables on emissions and the degree of uncertainty around those effects. To address this data limitation, the EPA decided to allow the EEMs to learn about effects of meteorological and farm-based predictor variables on NH_3 emissions simultaneously from swine and dairy.

In previous sections of this document, the terms "parameter" and "estimate" were used to describe the data and data collection methods used in the NAEMS. In this section, these terms are used in their formal statistical context. The term "parameter" refers to unknown constants (e.g., regression coefficients, variance components) whose values relate predictor variables to emissions and give the EEM its shape. In this section, parameters are designated using Greek letters. The term "estimate" refers to the best approximation of a parameter value by fitting the EEM to the NAEMS data. The estimate of a given parameter is designated using a "hat" symbol over the Greek letter. For example, the estimate for the intercept, β_0, would be denoted as $\widehat{\beta_0}$. The term "predict" refers to obtaining a value of emissions using the EEM, including the use of predictor variables and estimated parameters. Finally, the term "effect" does not imply a cause-effect relationship: rather, the phrase "the effect of *sa* on NH_3 emissions" refers to the regression coefficient applied to the variable *sa*.

Figure 5-1 shows the protocol that the EPA followed in developing the EEM. Phase 1 is the selection of the datasets to be used in EEM development. Dataset selection was based on the predictor variables that were monitored under the NAEMS as well as other important factors that affect emissions such as lagoon surface area. The dataset selection was also based on analyses of data completeness.

Phases 2 through 6 involve development and validation of the mathematical form of the EEM. The EEM has three components: the probability distribution, the mean trend function, and the covariance function. Equation 5-1 provides the general form of the EEM, and Table 5-1 defines the symbols and terms used in the equation.

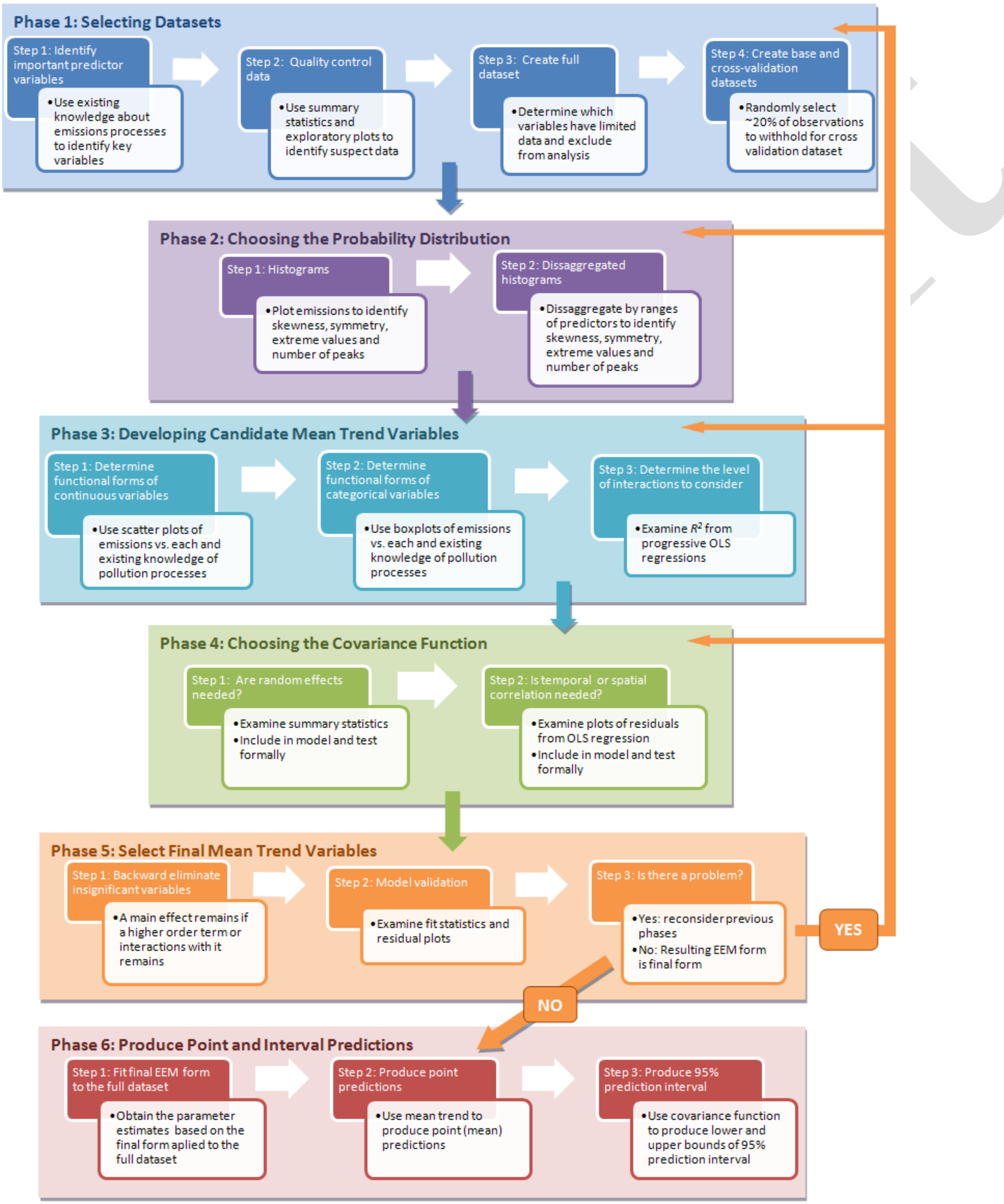


Figure 5-1. EEM Development Protocol

Table 5-1. Summary of Symbols and Terms Used in Equation 5-1

Description	Symbol
Index for sites	s
Index for dates	t
Mass of pollutant emitted from lagoon at site *s* on date *t*	Y_{st}
Index for regression coefficients, mean trend variables, and mean trend terms.	p
Number of mean trend variables, number of regression coefficients, number of mean terms minus one (the intercept is also a mean term).	P
Value of mean trend variable p for lagoon at site *s* on date *t*	x_{pst}
Regression coefficients	$\beta_p, \quad p = 1, \ldots, P$
Intercept	β_0
Mean trend terms	β_0 and $\beta_p\, x_{pht}$, $p = 1, \ldots, P$
Mean trend function— Three ways to write it.	$g(\mu_{st}) =$ $\beta_0 + \beta_1 x_{1st} + \cdots + \beta_p x_{pst} + \cdots + \beta_P x_{Pst}$ $= \boldsymbol{x}^T \boldsymbol{\beta}$
The $(P + 1) \times 1$vector containing a place-holder for the intercept plus the mean trend variables	$\boldsymbol{x} = (1 \quad x_{1st} \quad \ldots \quad x_{pst} \quad \ldots \quad x_{Pst})^T$
The $(P + 1) \times 1$vector containing the intercept and the regression coefficients	$\boldsymbol{\beta} = (\beta \quad \beta_1 \quad \ldots \quad \beta_p \quad \ldots \quad \beta_P)^T$
Deviation of emissions from lagoon at site *s* on date *t* from the value given by the mean trend function. A random variable. Optional for the gamma distribution.	e_{st}

$$\boldsymbol{Y_{st} \sim Gamma\,\{\mu_{st}, \phi\};\quad s = 1, \ldots, 9;\ \ t \text{ varies by site}}$$

$$\boldsymbol{g(\mu_{st}) = \beta_0 + \beta_1 x_{1st} + \cdots + \beta_p x_{pst} + \cdots + \beta_P x_{Pst}} \underbrace{\boldsymbol{+\, e_{st}}}_{\textbf{optional}} \qquad \text{Equation 5-1}$$

In the first line of Equation 5-1, the symbol Y_{st} represents pollutant emissions from the lagoon at site s and time t. The index, s, takes values 1 through 9, corresponding to the nine NAEMS sites. The index t identifies the date and time (in half-hour increments) associated with each half-hour NH_3 emissions value. Values of t are nested within values of s so that dates and times for different sites can be the same or different. Due to the rotating sampling schedule at 7 of the nine sites and the missing data, the dates were not always consecutive. Additionally, the half-hour values were not always consecutive for a given date.

The notation $Y_{st} \sim Gamma\ \{\mu_{st}, \phi\}$ says that NH_3 emissions at site s and time t follow the asymmetric gamma probability distribution. The mean of the distribution is represented by the parameter μ_{st} and the scale parameter, ϕ, that also determines the shape of the distribution. The prediction of NH_3 emissions for a given site and time is represented by $\widehat{Y_{st}}$ which is equal to the estimated mean, $\widehat{\mu_{st}}$.

In the second line of Equation 5-1, the expression $g(\mu_{st}) = \beta_0 + \beta_1 x_{1st} + \cdots + \beta_p x_{pst} + \cdots + \beta_P x_{Pst}$ says that the mean of the gamma distribution, μ_{st}, is related to the mean trend function, $\beta_0 + \beta_1 x_{1st} + \cdots + \beta_p x_{pst} + \cdots + \beta_P x_{Pst}$ by the link function, $g(\ )$. The mean trend function describes the relationship between the predictor variables and the expected value of NH_3 emissions. In the mean trend function, x_{pst} represents the value of the pth mean trend variable at site s and time t. The symbol β_p denotes the regression coefficient for that variable and the symbol β_0 represents the intercept. The mean trend variables differ from the predictor variables in that they represent the functional form through which the predictor variables enter the EEM. Lower-case p is an index for regression coefficients β_p, mean trend variables, x_{pst}, and their products (the non-intercept mean trend terms, $\beta_p x_{pst}$). The index, p, takes values 1,⋯, P, so that upper-case P is the number of non-intercept mean trend terms. Because the formula for the mean trend function is cumbersome, the matrix multiplication version of the function, $x^T\beta$, is used in the remainder of this section. The term $\boldsymbol{x}$ represents the $(P + 1) \times 1$ vector containing a placeholder for the intercept plus the mean trend variables, $x = (1 \quad x_{1st} \ \dots \ x_{pst} \ \dots \ x_{Pst})^T$, and the term $\boldsymbol{\beta}$ represents the $(P + 1) \times 1$ vector containing the intercept and the regression coefficients, $\boldsymbol{\beta} = (\beta \quad \beta_1 \ \dots \ \beta_p \ \dots \ \beta_P)^T$.

Link functions allow the flexibility to compare the fits of EEMs with different mathematical forms to the NH_3 emissions data. For the gamma distribution, there are three commonly used link functions: the identity, the natural logarithm and the reciprocal. When the identity link function is used (i.e., $g(\mu_{st}) = \mu_{st}$), the mean of the distribution is exactly equal to the mean trend function. When the natural logarithm or reciprocal link functions are used, the natural logarithm or the reciprocal of the mean of the distribution is equal to the mean trend function. Link functions are discussed in more detail in Section 5.4.1.

The optional term, e_{st}, represents deviations from the mean trend function is needed when there is covariance (or correlation) of some kind between values of NH_3 emissions.

The choice of probability distribution, the variables included in the mean trend function, the link function and the form of the covariance function were all based on analyses of a subset of the NAEMS data called the “base” dataset, validated using another subset called the “cross-validation” dataset, and then modified and re-validated when necessary. After the final mathematical forms were chosen, the EPA re-estimated the parameters using the “full” dataset (i.e., the combined base and cross-validation datasets).

The remainder of this section is organized as follows. Section 5.1 presents the selection of the candidate predictor variables and the base and cross-validation datasets. Section 5.2 explains why the gamma probability distribution was selected. Section 5.3 explains creation of candidate mean trend variables as functions of predictor variables. Section 5.4 discusses selection of the covariance function. Section 5.5 compares the fit statistics of three candidate EEMs with different mean trend variables. Section 5.6 describes how the EEMs are used to generate point and interval predictions for a half hour period and for sums of multiple half-hour periods. Section 5.6 also provides example predictions for each of the three EEMs discussed in Section 5.5. Section 5.7 discusses the challenges faced when developing the methodologies and offers additional analyses that may be considered by the EPA.

5.1 Selecting Datasets and Candidate Predictor Variables

The emissions and farm and lagoon data used to develop the NH_3 EEMs for lagoons were collected under the NAEMS from six swine sites (IA3A, IN4A, NC3A, NC4A, OK3A and OK4A) and three dairy sites (IN5A, WA5A and WI5A). The final reports submitted to the EPA contained daily lagoon emissions calculated based on measurements obtained using both the RPM model and the bLS model. In developing the NH_3 EEMs, the EPA used the measurements obtained using the RPM model. The EPA used the RPM data because these measurements were obtained using instrumentation and procedures that were similar to EPA’s developmental test method OTM-10 (Optical Remote Sensing for Emission Characterization from Non-Point Sources). The EPA did not use the bLS emissions measurements because these data were collected under the NAEMS to conduct a validation study of the bLS model performance relative to the RPM model. Furthermore, because the RPM emissions dataset is much larger than the bLS dataset, including the bLS measurements in the EEM development dataset would not provide any additional information on lagoon emissions. Therefore, the EPA chose to only use the RPM emissions data as it represented the largest unique NH_3 emissions dataset available from the NAEMS.

Due to the very limited number of daily NH_3 emissions values reported, the EPA prepared a database of half-hour values by combining the 30-minute emissions data and 5-minute meteorological data provided by the NAEMS researchers. To obtain values for each meteorological variable for a given half-hour, the EPA averaged the 5-minute data values within that half hour, as long as at least one of the values was not missing. Because it is unlikely that the value of any of these meteorological variables would vary enough within the consecutive 5-minute periods that comprise a half hour to significantly change the mean value of the variable during the half-hour, applying a completeness criterion to this aggregation would unnecessarily reduce an already limited data set. In the data aggregation, the EPA did not include any 30- or 5-minute data that had been flagged by the NAEMS researchers as invalid.

The resulting dataset contained a small number of negative NH_3 emissions (approximately 0.6 percent of the total values). Because such a small number of observations relative to the full dataset would be unlikely to provide information about variability in NH_3 emissions measurements, the EPA omitted these values from the EEM development dataset. Table 5-2 gives the number of 30-minute NH_3 emissions observations from each site after aggregation of the data, elimination of negative values, and evaluating completeness with meteorological variables. Table 5-3 shows the number of 30-minute values for the variables that were continuously monitored at each site after the EPA aggregated the 5-minute data to 30 minute values.

Table 5-2. Number of 30-Minute NH_3 Emissions Values by Site

Monitored Site		Site Number in EEM	No. of NH_3 Emissions Values
Swine breeding and gestation	IN4A	1	677
	NC4A	2	1,193
	OK4A	3	1,780
Swine grow/finishing	IA3A	4	1,048
	NC3A	5	1,074
	OK3A	6	1,600
Dairy	IN5A	7	2,593
	WA5A	8	507
	WI5A	9	312
All sites			10,784

Table 5-3. Number of 30-Minute Data Values for Continuous Variables

Description	Units	Swine Breeding and Gestation Sites			Swine Grow/Finishing Sites			Dairy Sites			All Sites
		1	2	3	4	5	6	7	8	9	
		IN4A	NC4A	OK4A	IA3A	NC3A	OK3A	IN5A	WA5A	WI5A	
Ambient Meteorological Data											
Ambient temperature	°C	677	1,193	1,780	1,048	1,078	1,600	2,593	507	312	10,788
Relative humidity	%	698	1,253	2,319	1,059	1,122	1,618	3,683	529	350	12,631
Atmospheric pressure	kPa	678	1,194	1,683	1,048	957	1,587	2,669	520	300	10,636
Dew point temperature	°C	677	1,193	1,780	1,048	1,074	1,600	2,593	507	312	10,784
Solar radiation	W/m^2	522	924	2,007	841	991	1,297	2,837	440	293	10,152
Surface wetness	Resistance (Ω)	359	520	700	359	301	537	1,298	29	73	4,176
Wind speed	m/s	711	1,253	2,324	1,059	1,126	1,795	3,683	551	352	12,854
Lagoon Liquid Data											
pH	pH	357	1,191	1,836	0	983	1,211	1,330	0	12	6,920
Oxidation/ reduction potential	Millivolts (mV)	543	1,238	2,080	0	931	1,573	1,396	0	33	7,794
Temperature	°C	581	1,237	2,077	0	888	1,571	1,422	0	73	7,849

5.1.1 Candidate predictor variables

In addition to the continuous variables shown in Table 5-3, the EPA compiled the following information from the final reports for each site: type and presence of natural cover on the lagoon surface, lagoon liquid composition and animal inventory. This information was collected intermittently for select days during the NAEMS (see Section 4). For EEM development, the EPA assumed that the value of the intermittent data was constant for each 30-minute period of the day in which the information was collected. Table 5-4 shows the number of 30-minute values available for the intermittent data based on the EPA's approach.

To supplement the continuous and intermittent data, the EPA compiled information regarding the farms and lagoons at each monitoring site from the NAEMS open-source QAPP, SMPs and final reports. Table 5-5 shows the farm and lagoon information and whether the information was available for each monitoring site. The information listed in Table 5-5 consists of categorical specifications (e.g., animal type, impoundment type) where the value is constant and not expected to vary over time, and discrete specifications (e.g., animal weight, lagoon liquid depth) where the reported value is constant but the actual value was expected to change over time. As with the intermittent data, the EPA assumed that the values of the farm and lagoon information were constant throughout the NAEMS sampling period Therefore, the values were applied to all 30-minute periods.

In determining which of the data and information shown in Table 5-3, Table 5-4 and Table 5-5 would be selected as candidate predictor variables for EEM development, the EPA's primary consideration was data completeness. However, the EPA also considered whether or not the variable would be readily available to farmers. Table 5-6 shows the candidate predictor variables considered in the development of EEMs for lagoon NH_3 emissions.

Table 5-4. Number of 30-Minute Values Available for Intermittent Data by Site

Monitoring Site		Lagoon Cover	Lagoon Liquid Composition			Animal Inventory
			Nitrogen Content	Solids Content	NH_3 Content	
Swine Breeding and Gestation	IN4A	83	0	0	0	0
	NC4A	97	48	0	0	0
	OK4A	149	48	0	0	2,234
Swine grow/finishing	IA3A	122	96	144	144	1,059
	NC3A	31	144	0	0	1,126
	OK3A	136	96	96	96	1,795
Dairy	IN5A	300	0	0	0	3,683
	WA5A	11	0	0	0	551
	WI5A	24	0	0	0	352
All sites		953	432	240	240	10,890

Table 5-5. Farm and Lagoon Information by Site

Variable Description		Units	Variable Type[a]	Swine Breeding and Gestation Sites			Swine Grow/Finishing Sites			Dairy Sites		
				IN4A	NC4A	OK4A	IA3A	NC3A	OK3A	IN5A	WA5A	WI5A
Lagoon site ID	NAEMS site ID.	NA[b]	Cat.	Y[c]	Y	Y	Y	Y	Y	Y	Y	Y
Animal type	Swine or dairy.	NA	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Farm animal capacity	Average animal inventory contributing manure to the monitored lagoon.	head	Cont.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Average animal weight	Average weight of animals at swine and dairy farms.	kg	Cont.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Average animal weight (piglet)	Average weight of piglets at swine breeding and gestation farms.	kg	Cont.	Y	N	N	NA	NA	NA	NA	NA	NA
Number of confinement structures on the farm	Design specification for the number of confinement structures that contributed manure to the lagoon.	structures	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Farm manure management system	Design specification for the type of system (e.g., pull-plug pit) used to remove manure from the confinement structure.	NA	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Solids separation	Design specification to indicate whether the site separates solids from manure before it is discharged to the lagoon.	NA	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Odor control	Design specification to indicate whether the site used a lagoon additive that suppresses NH_3 emissions.	NA	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Farm age	Date of farm construction (or since major renovation)	years	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 5-5. Farm and Lagoon Information by Site

Variable Description		Units	Variable Type[a]	Swine Breeding and Gestation Sites			Swine Grow/Finishing Sites			Dairy Sites		
				IN4A	NC4A	OK4A	IA3A	NC3A	OK3A	IN5A	WA5A	WI5A
Impoundment type	Lagoon or basin.	NA	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lagoon configuration	Design specification for the lagoon configuration (e.g., single stage, multiple stage).	NA	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lagoon volumetric loading rate	Design specification for the volumetric loading rate of organic material, defined as the mass of volatile solids (VS), added to the lagoon daily.	lb VS/d-1,000 ft^3	Cat.	Y	Y	Y	N	Y	N	N	N	Y
Lagoon surface loading rate	Design specification for the surface loading rate of organic material added to the lagoon each day.	lb VS/d-ac	Cat.	Y	Y	Y	N	Y	N	N	N	N
Lagoon volume	Design specification for the volumetric capacity of the lagoon to store liquid and accumulated sludge.	ft^3	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lagoon surface area	Design specification for the total surface area of the lagoon.	1,000 ft^2	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lagoon liquid depth	Design specification for the average depth of the liquid above the layer of accumulated sludge at the bottom of the lagoon.	ft	Cat.	Y	Y	Y	Y	N	Y	Y	Y	Y
Lagoon sludge depth	Design specification for the depth of accumulated sludge at the bottom of the lagoon.	ft	Cat.	N	Y	N	Y	Y	Y	Y	N	Y
Number of manure inlets to the lagoon	Design specification for the number of inlets that discharge manure to the lagoon.	inlets	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 5-5. Farm and Lagoon Information by Site

Variable Description		Units	Variable Type[a]	Swine Breeding and Gestation Sites			Swine Grow/Finishing Sites			Dairy Sites		
				IN4A	NC4A	OK4A	IA3A	NC3A	OK3A	IN5A	WA5A	WI5A
Manure discharge schedule	Design specification for the frequency of manure removal from confinement structures	days	Cat.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lagoon pump-out frequency	Design specification that indicates the frequency that liquid is removed from the lagoon.	days	Cat.	Y	N	Y	Y	N	Y	N	N	Y
Lagoon agitation prior to pump-out	Activity indicator to note that the farm agitated the lagoon contents prior to pump-out.	1 = activity 0 = no activity	Cat.	Y	Y	Y	Y	Y	Y	N	N	N
Manure discharge to lagoon event	Activity indicator for a given date to note when manure was removed from the confinement structure	1 = varies[d] 0 = no activity	Cat.	Y	Y	N	Y	N	N	N	N	Y

[a] Cont. = Continuous, the value of the variable can change over time. Cat. = Categorical, the value of the variable is constant over time. Disc. = Discrete, the value is constant, but can change over time.
[b] NA = Not applicable.
[c] Y = data is available, N = data is not available.
[d] 0 = no removal; 1 = pit agitation and pump out; 2 = drained/pumped; 3 = scrape; 4 = pump to solids separation, excess to lagoon; blank = not reported.

Table 5-6. Selected Candidate Predictor Variables

Category	Description	Predictor Variable	Units
Ambient	Temperature	*ta*	^{o}C
	Relative humidity	*ha*	%
	Wind Speed	*ws*	m/s
	Hour of the day	*hour*	hour
	Julian day (day of the year)	*jday*	day
Farm and lagoon	Animal type (Dairy or Swine indicator)	*animal*	NA[a]
	Capacity of farm (number of animals)	*capacity*	head
	Average adult animal weight	*adultwt*	lb
	Number of confinement structures	*barns*	barns
	Manure management system	*mms*	NA
	Surface area	*sa*	1000 ft^2
	Number of manure inlets into lagoon	*inlets*	inlets
	Whether an odor control agent was used on a given day	*odorctrl*	NA

[a] NA = Not applicable.

5.1.1.1 *Ambient Meteorology*

The NH_3 emissions rate from lagoons is affected by air temperature, wind speed across the lagoon surface and relative humidity (NAS, 2002). The ambient air temperature determines the temperature of the lagoon liquid, particularly near the surface of the lagoon. With increasing ambient air temperature, the resulting increase in liquid temperature is expected to increase NH_3 emissions. This positive relationship between liquid temperature and NH_3 emissions from swine lagoons was modeled by Aneja *et al.* (2001) and Harper *et al.* (1998). Values for ambient temperature were available for all swine and dairy sites. Consequently, the EPA included ambient temperature as a candidate predictor variable.

With regard to wind speed, the EPA expected that, in general, higher wind speeds would increase NH_3 emissions. As the wind sweeps across the lagoon surface, the concentration gradient between the lagoon liquid and the liquid-air boundary layer at the surface decreases, thereby increasing diffusion of NH_3 to the atmosphere. Also, high winds can increase surface agitation, thereby effectively increasing the surface area available for diffusion of NH_3 to the atmosphere. The positive relationship of wind speed on NH_3 emissions was noted in the model for swine lagoons developed by Harper *et al.* (1998). Values for wind speed were available for all swine and dairy sites. Consequently, the EPA included wind speed as a candidate predictor variable.

The EPA expected that increases in relative humidity above the lagoon surface would tend to decrease the diffusion of gas molecules from the liquid/gas interface. Values for ambient relative humidity were available for all swine and dairy sites. Therefore, the EPA included relative humidity as a candidate predictor variable.

The EPA also expected that increases in atmospheric pressure above the lagoon surface would tend to decrease NH_3 emissions. Higher atmospheric pressure will decrease the gradient between the partial pressure of NH_3 gas dissolved in the lagoon liquid and the atmosphere above the lagoon surface, thereby reducing diffusion of gas molecules from the liquid. However, the EPA did not use atmospheric pressure as a predictor variable because to do so would have significantly reduced the size of the dataset.

The dew point is the temperature at which moisture in the atmosphere condenses to liquid water. This condensed moisture could reabsorb NH_3 released from the lagoon and potentially reduce NH_3 emissions. The condensed water could also reduce NH_3 emissions by reducing the concentration gradient between NH_3 in the bulk lagoon liquid and the boundary layer near the lagoon surface. Dew point is a function of temperature and humidity. Based on exploratory data plots, the EPA determined that the interaction effect of ambient temperature and relative humidity would be more effectively captured using the product of ambient temperature and relative humidity, rather than dew point values. Consequently, the EPA did not include dew point temperature as a candidate predictor variable.

Because rainfall incident to the lagoon could dilute the concentration of NH_3 near the lagoon surface, the EPA expected that wet conditions would result in lower NH_3 emissions. An electrical resistance grid was used to indicate the presence of wet or dry conditions on the ground near the lagoons; however, the amount of precipitation received by each site was not collected under the NAEMS. The number of wetness data values available for all sites was very limited (4,176 30-minute values for all sites) and incorporation of this data would reduce the size of the dataset available to develop the EEMs by approximately 67 percent. Consequently, the EPA did not include data for wet/dry conditions as a candidate predictor variable.

Because solar radiation affects ambient air temperature and the temperature of the liquid on the surface of the lagoon, increases in solar radiation were expected to increase NH_3 emissions. However, the EPA chose not to use solar radiation as a predictor variable because this information is not expected to be readily available. While temperature, humidity, and wind speed are recorded routinely at a network of meteorological sites, solar radiation is not. To capture the effect of the annual solar radiation cycle on NH_3 emissions, the EPA used the Julian day (*jday*) corresponding to the measurement time stamps as a predictor variable. The Julian day is the day of the year, so that *jday* can take values 1 through 365 or 366, depending on whether or not the day falls in a leap year. The EPA believed that interactions of *jday* with *ta* and *ha* should account for reduced solar radiation due to cloud cover. Also, to account for the effect of the diurnal solar radiation cycle, the EPA used a function of *hour*.

5.1.1.2 ***Lagoon Liquid Data***

The NAEMS collected continuous monitoring data for pH, oxidation/reduction potential (ORP), and liquid temperature and intermittent data for the nitrogen, ammoniacal nitrogen and solids content of the liquid.

The concentration of NH_3 at the lagoon surface depends on the total concentration of ammoniacal nitrogen [i.e., NH_3 plus ammonium ion (NH_4^+)] in the liquid. The balance between NH_3 gas and highly-soluble NH_4^+ is a function of pH and temperature of the lagoon liquid. At a constant pH, the concentration of NH_3 relative to NH_4^+ increases, thereby increasing NH_3 emissions. At a constant temperature, the concentration of NH_3 relative to NH_4^+ also increases, thereby increasing lagoon NH_3 emissions potential (NAS, 2002). As noted in Section 5.1.1.1, previous studies [Aneja *et al.* (2001) and Harper *et al.* (1998)] have shown a positive relationship between liquid temperature and NH_3 emissions from swine lagoons. Consequently, the EPA expected that increasing pH or lagoon liquid temperature or both would result in higher NH_3 emissions.

The ORP measures the activity of electrons in a liquid and indicates whether the liquid is an oxidizing or reducing environment. However, because the mineralization of organic nitrogen to NH_3 occurs in both oxidizing and reducing environments, the ORP value only indicates if the potential exists for the reduction of NH_3 emissions due to nitrification. Therefore, the EPA expected that NH_3 emissions would decrease with positive ORP values indicating an aerobic and oxidizing environment.

Although the pH, ORP and temperature of the lagoon liquid were expected to affect NH_3 emissions from lagoons, the EPA did not include these data as candidate predictor variables due to the limited number of data values and because data were not provided for sites IA3A and WA5A.

Because the organic and ammoniacal nitrogen present in the lagoon liquid are precursors to NH_3 emissions, the EPA expected that NH_3 emissions would be higher at lagoons with higher total nitrogen concentrations. Nitrogen compounds can be bound to lagoon solids thereby preventing the release of ammonia precursors into the bulk lagoon liquid; therefore, the EPA also considered using the solids content of lagoon liquid in EEM development. However, the EPA did not include data for the nitrogen or NH_3 content of the lagoon liquid as candidate predictor variables due to the limited number of data values (see Table 5-4).

At higher volumetric loading rates, the concentrations of total solids will be higher as will the concentration of total nitrogen, which is comprised of organic and ammoniacal nitrogen. Thus, the potential for NH_3 emissions also will be higher because the concentration of free NH_3 is a function of the ammoniacal nitrogen concentration with the latter a function of the readily mineralizable fraction of organic nitrogen. However, the EPA did not include data for total solids as candidate predictor variables due to the limited number of data values (see Table 5-4).

5.1.1.3 ***Farm and Lagoon Information***

For animal type, the EPA included an indicator variable to designate whether the lagoon was located at a swine or dairy farm as a candidate predictor variable to consider the possible effect on lagoon NH_3 emissions due to the different type of animal. The EPA did not include a variable to distinguish between swine breeding/gestation and grow/finish farms. Although the swine industry includes farms that have breeding/gestation and grow/finish production stages on the same site, the NAEMS did not measure emissions at any farrow-to-finish operations. Consequently, including a variable to distinguish farm type in the EEM development dataset would have limited the applicability of the EEM to either a breeding/gestation or a grow/finish farm.

Because organic nitrogen in excreted manure is the precursor to NH_3 emissions from the lagoon, farms that generate more manure were expected to have larger lagoons and subsequently higher NH_3 mass emissions. To account for the relative size of the farms at each monitoring site, the EPA considered using the farm animal capacity, average daily animal inventory, average animal weight and the number of confinement structures as candidate predictor variables. Values for animal capacity, average animal weight and the number of confinement structures were available from the NAEMS documentation for each site; therefore, these data were used as candidate predictor variables. However, daily animal inventory values were not reported for sites IN4A and NC4A. Consequently, the EPA did not include animal inventory as a candidate predictor variable. In addition, the average animal weight for growing/finishing houses was the average weight over the lifetime of the pig, and therefore fall between the initial weight of the piglet (approximately 14 lb) and the final market weight of the pig (approximately 247 lb).

The swine and dairy sites used a variety of methods (e.g., pull-plug pit, flushing, scraping) to remove manure from the confinement houses. The EPA included the type of manure collection system used at each site as a candidate predictor variable to determine whether the manner in which manure was removed from the houses was related to NH_3 emissions from lagoons.

Dairies typically use some form of solids separation step to remove solids and inorganic material (e.g., sand used for bedding) prior to discharging manure to the lagoon and all of the NAEMS dairy sites use solids separation. Conversely, none of the swine sites used solids separation which is typical of swine operations due to the consistency of excreted manure. Because the presence of solids separation at the sites does not vary within animal type, the effect on lagoon NH_3 emissions attributable to the presence or absence of solids separation cannot be assessed using this information. Consequently, the EPA did not include these data as a predictor variable.

Because increased mixing of the lagoon liquid could increase NH_3 emissions due to exposure of more concentrated liquid to the lagoon surface, the EPA considered including the following information to characterize the potential for turbulence in the lagoon in the list of candidate predictor variables: number of lagoon manure inlets, manure discharge schedule and lagoon liquid pump-out frequency. The EPA also considered using an indicator variable to designate days when the lagoon liquid was agitated prior to a pump-out event. Data values for the manure discharge schedule and the number of lagoon inlets were available for all sites; therefore, the EPA included these data as candidate predictor variables. The EPA did not include the frequency with which liquid was pumped out of the lagoon or indicator variables to denote specific events in the list of candidate predictor variables because these data were not reported for all sites.

The EPA considered using the following design specifications as predictor variables to assess whether lagoon design was related to NH_3 emissions: impoundment type (i.e., lagoon or basin), configuration, loading rates, volume, surface area, liquid depth and sludge depth. However, the EPA did not include lagoon loading rates, liquid depth and sludge depth because these data were not available for all sites. Also, the EPA did not consider using a predictor variable for single- or multiple-stage lagoons. Only site WI5A had a multi-stage lagoon, based on the descriptions provided in the SMP and final report. However, the emissions from each of the three stages were not measured independently. The monitoring equipment was located such that the total emissions from stages 1 and 2 were measured. Emissions from the 3rd stage, which was used to supply flush water, were not measured under the NAEMS. Consequently, the EPA decided to use the total emissions and total surface area from stages 1 and 2 as representative of a single-stage lagoon, rather than exclude site WI5A due to the different lagoon configuration. Therefore, a predictor variable for lagoon stages was not used.

The presence of a natural cover (i.e., crust, scum or ice) on the lagoon surface tends to reduce emissions because the cover inhibits diffusion of NH_3 from the lagoon liquid to the atmosphere. Observations regarding the type (e.g., crust, ice) and degree of cover (percent of surface area) were provided for select days at each site. Although the presence of natural lagoon cover (i.e., crust, scum or ice) was expected to affect NH_3 emissions, the EPA did not use it as a candidate predictor variable because of the limited number of recorded observations.

One of the sites reportedly used an additive that reduces odor emissions from the lagoon. Consequently, the EPA included a predictor variable to note whether the grower used an odor control additive in the lagoon.

The EPA also did not use the type of impoundment (i.e., lagoon or basin), as a predictor variable. For impoundment type, the QAPP, SMPs and the final reports do not define the design and operational differences between a "lagoon" and "basin" and the documents tend to use the terms interchangeably. Based on discussions with the NAEMS Science Advisor, dairies tend to use basins which have a lower degree of microbial activity than lagoons. Using this information, all of the dairy sites would be assigned a basin for impoundment type and all of the swine sites would be assigned a lagoon. However, because the type of impoundment at the sites does not vary within animal type, the effect on lagoon NH_3 emissions attributable to the type of impoundment cannot be assessed using this information. Consequently, the EPA did not include impoundment type as a predictor variable.

5.1.2 Full, Base and Cross-validation Datasets

To ensure that the data selected for EEM development were representative of more than one of the monitored sites, the EPA limited the dataset for use in EEM development to those records for which data values were available for all selected predictor variables.

The number of 30-minute observations for which NH_3 emissions and all of the selected candidate predictor variables were available was 10,783. This dataset is referred to as the "full" dataset. From the full dataset, the EPA withheld a cross-validation dataset for use in testing the EEMs and ensuring that any relationships predicted by the EEMs would apply generally, rather than being attributable to anomalies that might be present in the data. To create the cross-validation dataset, the EPA randomly selected 2,191 (approximately 20 percent) of the 10,783 observations in the full dataset. The remaining 8,592 observations comprised the "base" dataset.

Ideally, 20 percent of the observations representing different combinations of conditions would be withheld to ensure that neither the cross-validation dataset nor the base dataset over or under-represents any one set of conditions. Because there are so many predictor variables for the NH_3 lagoon dataset, sub-dividing the data into many different sets of conditions would result in subsets with few observations, and it would be difficult to ensure that 20 percent of any one set is actually withheld.

To ensure that the cross-validation and base datasets selected initially did not disproportionately over- or under-represented one or more sets of conditions, the EPA created two additional cross-validation datasets and corresponding base datasets by randomly selecting approximately 20 percent of the observations to withhold. Although this report presents results for the first base and cross-validation datasets only, the EPA checked for notable aberrations in results for the other two cross-validation datasets as well.

5.2 Selecting the Probability Distribution

Identifying the appropriate probability distribution allows for production of accurate prediction intervals that quantify uncertainty in the point estimate of NH_3 emissions. Selection of the correct distribution also ensures the validity of the p-values used to determine the statistical significance of regression coefficient estimates.

To determine the probability distribution, the EPA produced a series of histograms of NH_3 emissions, beginning with a histogram of the full dataset. The histogram in Figure 5-2 shows that the empirical distribution of NH_3 emissions is unimodal and is skew right. Note that the empirical distribution is increasing and then decreasing, rather than strictly decreasing.

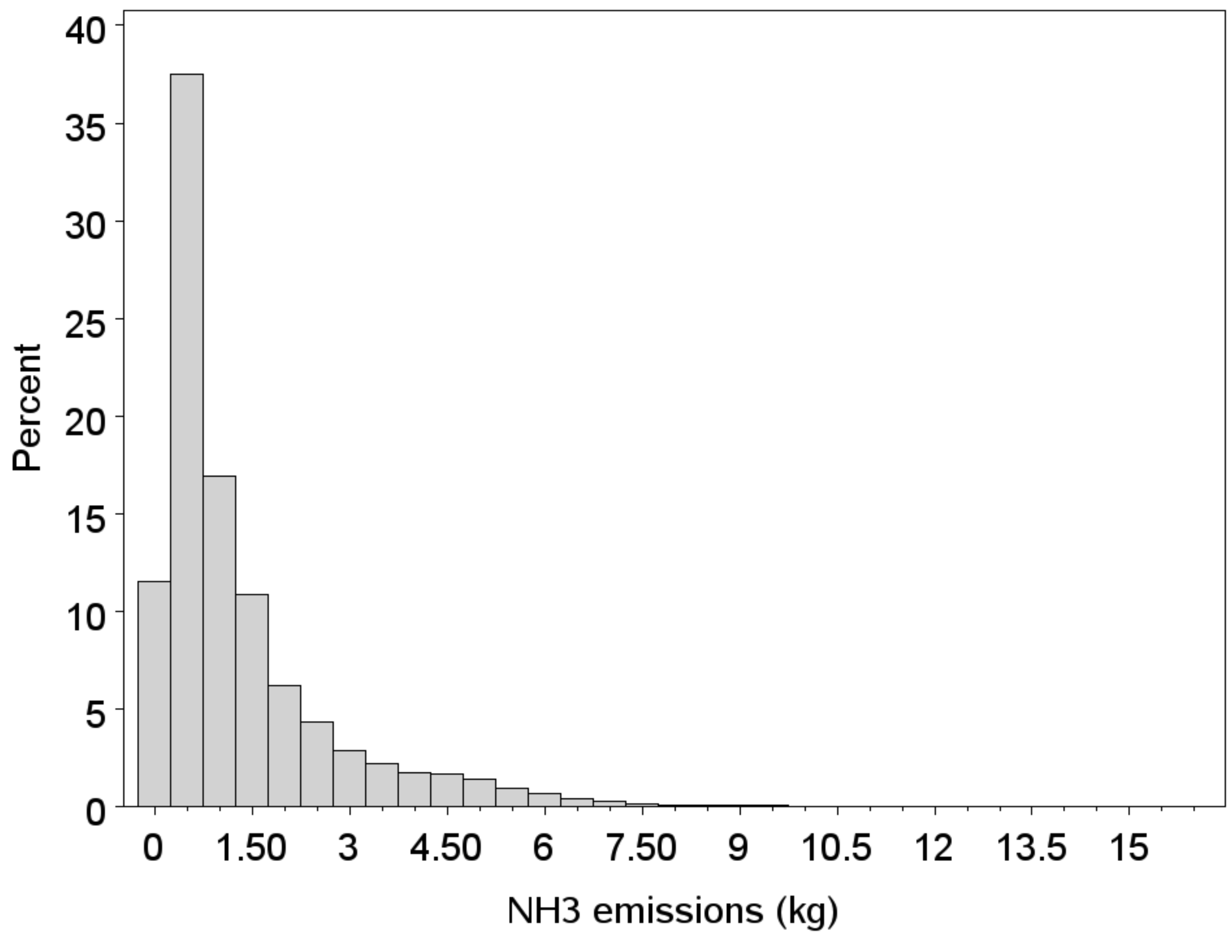


Figure 5-2. Histogram of NH_3 Emissions

This figure makes it appear that a skew-right distribution would be appropriate, but aggregating NH_3 emissions resulting from many different sets of conditions into a single histogram could mask differences in the distribution under different conditions, and could make a symmetric distribution appear skew. The EPA therefore disaggregated the dataset into bins based on values of the meteorological predictor variables. The EPA created three new variables (*tabin*, *habin* and *wsbin*) to identify eight ranges of values of ambient temperature, humidity and wind speed, respectively. These meteorological variables were used because they had a wide range of values that provided varying conditions to examine the distribution. These variables also displayed a significant ability to predict NH_3 emissions based on initial tests. These meteorological bins will also be used later in the analysis to allow for a closer examination of the data to verify that a wide range of meteorological conditions are covered by the dataset, as well as the emissions trends for combinations of these conditions (e.g., cold, moist and windy conditions versus hot, dry and calm conditions). The bin cut-off values for each are given in Table 5-7. In the first temperature bin, *tabin* = 1, for example, the values of temperature fall between -29 and -20 ^{o} C.

Table 5-7. Meteorological Variable Bin Cut-Offs

Ambient Temperature (° C)				Relative Humidity (%)				Wind Speed (m/s)		
tabin	**min**	**max**		***habin***	**min**	**max**		***wsbin***	**min**	**max**
1	-29	-20		1	0.0	12.5		1	0.00	2.25
2	-20	-11		2	12.5	25.0		2	2.25	4.50
3	-11	-2		3	25.0	37.5		3	4.5	6.75
4	-2	7		4	37.5	50.0		4	6.75	9.00
5	7	16		5	50.0	62.5		5	9.00	11.25
6	16	25		6	62.5	75.0		6	11.25	13.50
7	25	34		7	75.0	87.5		7	13.50	15.75
8	34	43		8	87.5	100.0		8	15.75	18.00

Figure 5-3 shows histograms of NH_3 emissions disaggregated by *tabin*. Within the eight different ranges of temperatures, bins 2 through 8 have enough data to draw conclusions. The distribution of NH_3 emissions is still primarily increasing then decreasing, unimodal and primarily skew right. Figure 5-4 shows a further disaggregation into combinations of *tabin* and *wsbin*. The temperature and wind speed were used for this disaggregation because preliminary plots showed that NH_3 emissions had the strongest relationship with these two variables. Because not all of the 64 combinations of *tabin* and *wsbin* contain many observations, the 25 combinations in Figure 5-4 are those with the most observations. Notice that the predominant shape of the distribution for these subsets of the data is again unimodal and skew right, although some plots are somewhat symmetric and others are skew left.

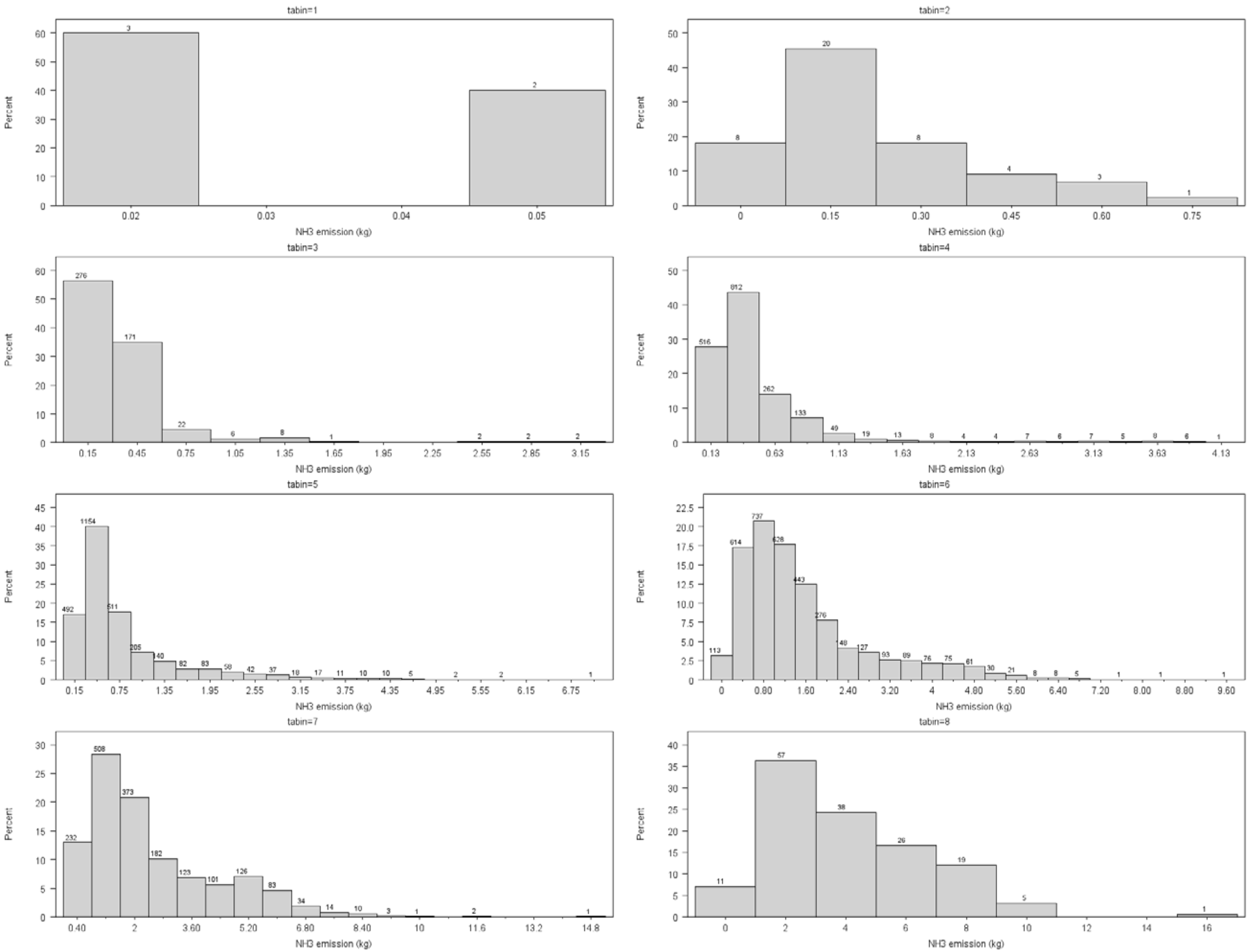


Figure 5-3. Histogram of NH_3 Emissions by Temperature Bin

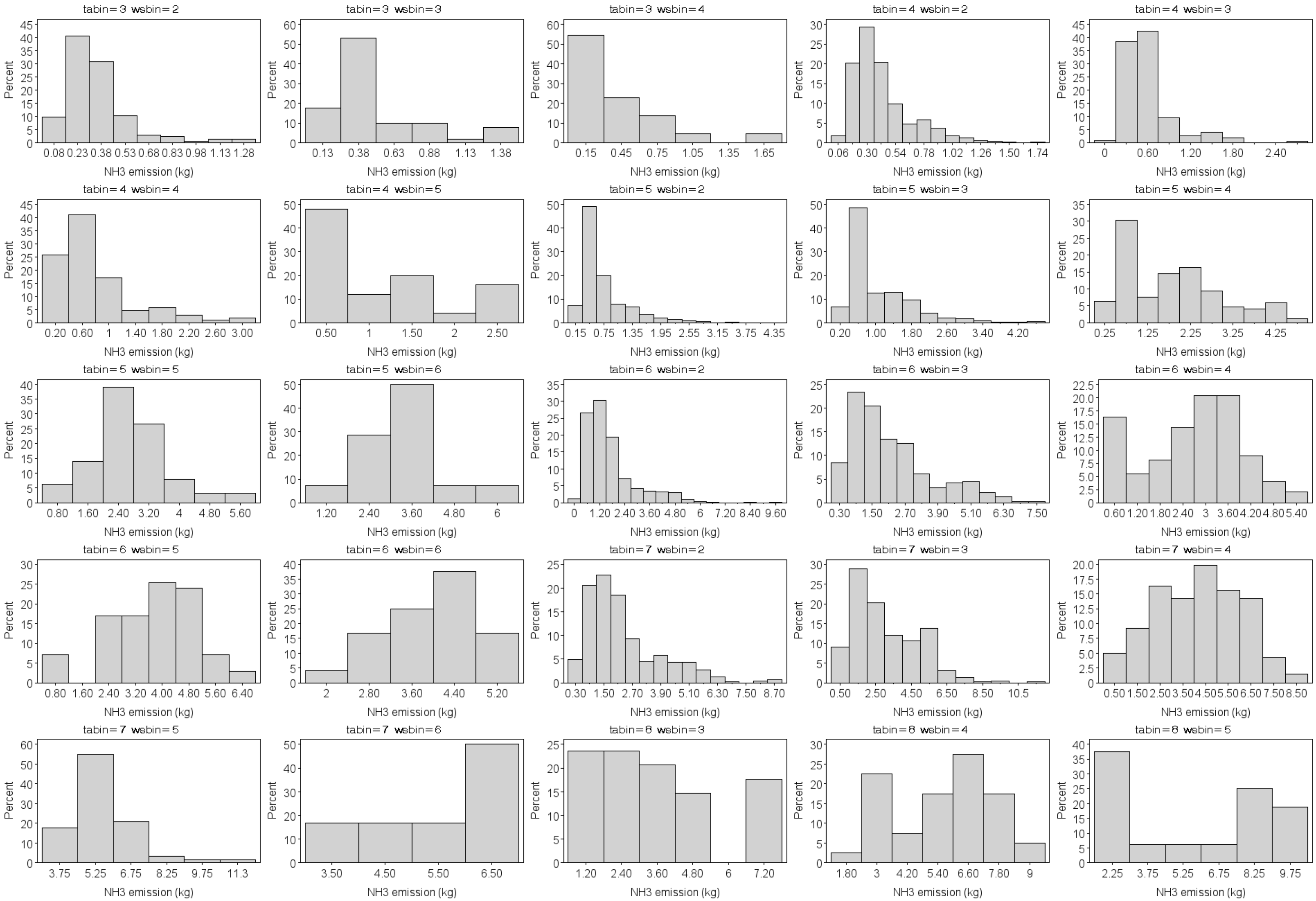


Figure 5-4. Histogram of NH_3 Emissions by Temperature and Wind Speed Bins

The EPA concluded that using a normal distribution for the NH_3 emissions was inappropriate for EEM development. To capture the skew-right nature of the NAEMS NH_3 emissions data, the EPA chose the gamma distribution. An advantage of using the gamma distribution as opposed to the often-used lognormal distribution is that use of the gamma distribution allows the flexibility of relating the mean of NH_3 emissions to predictor variables via a link function. The lognormal distribution relates mean NH_3 emissions to the mean trend function only by way of the natural log.

The gamma distribution is a skew-right distribution parameterized by the mean parameter, μ, and the parameter ϕ, which must be greater than 0 and describes the degree of skewness in the distribution. Figure 5-5 shows the different shapes the gamma distribution takes for different values of ϕ when $\mu = 1.5$, which is a typical value of NH_3 emissions from lagoons. When ϕ falls between 0 and 1, the distribution is increasing and then decreasing, unimodal and skew right, like most of the histograms in Figure 5-2, Figure 5-3 and Figure 5-4. When ϕ is greater than 1 (yellow line in Figure 5-5), the distribution is strictly decreasing, like a ski slope, and like some of the histograms in Figure 5-4. When ϕ approaches 0 (the blue line in Figure 5-5), the distribution is symmetric and bell-shaped, like some of the histograms in Figure 5-4.

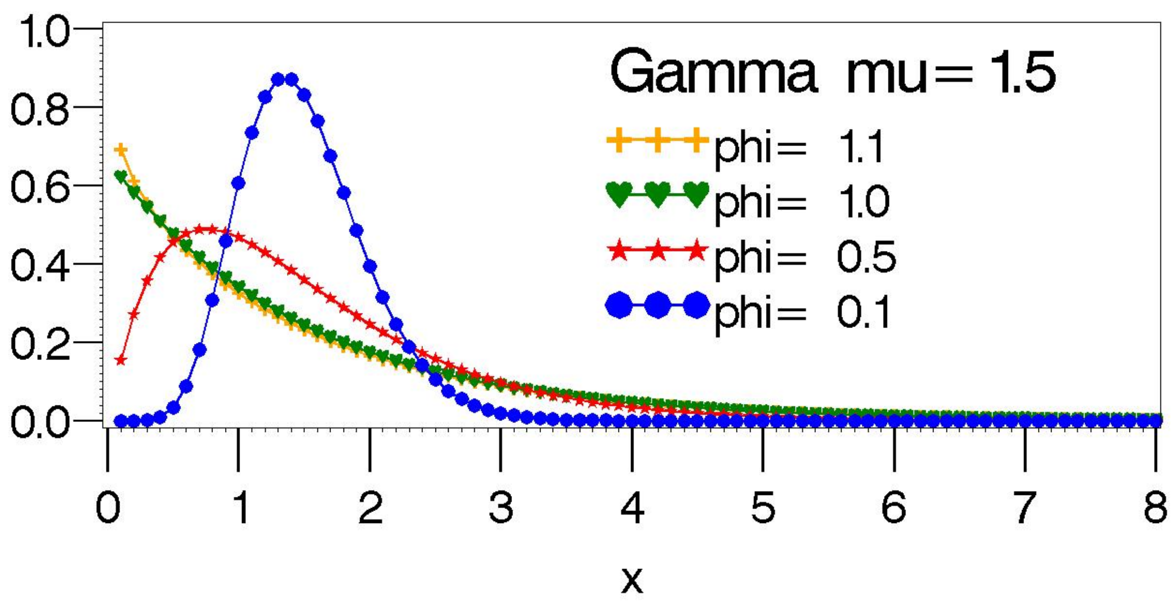


Figure 5-5. Shapes of Gamma Distribution for Different Parameter Values

5.3 Developing Candidate Mean Trend Variables

The mean trend function quantifies the relationship between the expected value of NH_3 emissions and the predictor variables. When combined with the link function, the mean trend function provides the formula that is used to obtain the point estimate. Section 5.1.1 discusses the challenges of deciding how many and which of the static farm-based predictor variables to include, and shows how unbalanced coverage of meteorological conditions from site-to-site presents an obstacle to using the NAEMS data to learn about the effects of any of the farm-based variables on NH_3 emissions. Sections 5.3.1 through 5.3.5 describe how functional forms were chosen to represent effects of the continuous predictor variables on NH_3 emissions. Section 5.3.7 explains the selection of a working mean trend function, based on the continuous predictors, which will be used in Section 5.4 to make decisions regarding the covariance and link functions.

Table 5-8 presents the summary statistics for NH_3 emissions and the ambient meteorological variables. These summary statistics provided the context within which decisions regarding the form of the mean trend function were made. The mean value of NH_3 emissions was 1.3 kg, with a standard deviation of 1.4 kg. The minimum value is 0 kg, and the maximum observed half-hour emissions value was 16 kg. Notice that the median, 0.77 kg, is less than the mean, which is always the case for skew-right distributions. The middle 50 percent of emissions fall between the lower and upper quartiles, 0.38 and 1.7 kg, respectively.

Table 5-8. Summary Statistics for NH_3 and Meteorological Variables

Variable[a]	Description	Mean	Std Dev	Min	Lower Quartile	Median	Upper Quartile	Max
nh_3	NH_3 emissions (kg)	1.30	1.42	0.00	0.38	0.77	1.66	16.09
ta	Temperature (^{o} C)	15.35	10.13	-23.29	8.33	16.40	22.94	41.37
ha	Relative humidity (%)	62.35	22.21	6.58	44.58	63.91	81.43	100.00
ws	Wind speed (m/s)	3.34	2.36	0.10	1.70	2.80	4.30	17.40

[a] Summary statistics for *hour* and *jday* are not included.

Ambient temperature in the NAEMS data ranges from -23 to 41 ^{o}C, with the middle 50 percent of values falling between 8 and 23 ^{o}C. The distribution of temperature is more symmetric than the distribution of NH_3 emissions, and if it is skew, it is skew to lower temperatures (meaning there are a few extreme low temperatures that would pull a tail to the left.) Notice that the mean is slightly less than the median.

Relative humidity can only take values from 0 to 100, and in this dataset it ranges from 6.6 to 100, with the middle 50 percent of values falling between 45 and 81 percent. The mean and median are close together, and the distribution is somewhat symmetric. Wind speed values range from 0 to 17 m/s. The distribution of wind speed is quite skew right; 75 percent of the values fall below 4.3 m/s, but there are a handful of extreme values such as 17 m/s.

Table 5-9 summarizes the candidate mean trend variables chosen as the functional forms through for each predictor variable in the EEM.

Table 5-9. Summary of Main Effect Mean Trend Variables

Main Effect	Original Predictor Variable(s)[a]	Functional Form Chosen	Centering Value	Scaling Value
animal	animal	Same as original variable	NA[b]	NA
size	capacity*	size* = (capacity*)(adultwt*), Linear	1,977,855	1,845,378
	adultwt*			
ta	ta*	Linear	16	10
ha	ha*	Linear	61	22
ws	ws*	Linear	3.2	2.2
sa	sa*	Linear	160	89
sh	Solar radiation via hour	As defined in Equation 5-2	NA	NA
ch				
sd	Solar radiation via jday	As defined in Equation 5-3	NA	NA
cd				

[a] An asterisk (*) is used to note that these predictor variables are the original values reported in the data submitted to the EPA before centering and scaling.
[b] NA = Not applicable.

5.3.1 Temperature

To determine the functional form through which the predictor variable for ambient temperature, *ta*, should enter the mean trend function, the EPA plotted ln(NH_3) versus *ta* for each animal type. The plots use ln(NH_3) on the y-axis rather than NH_3 because the log link function was chosen to represent the relationship between the mean of the gamma distribution and the mean trend function: $\ln(\mu_{st}) = x^T\beta$. Because the log function is a strictly increasing function, if ln(NH_3) is an increasing function of any predictor variable, then so is NH_3. Therefore conclusions based on ln(NH_3) emissions are directly applicable to NH_3 emission trends. Section 5.4.1 shows why this decision was made and makes it clear that the EEM development process involves a nonlinear decision-making process, as illustrated in Figure 5-1, in which decisions are re-evaluated based on results of the EEM validation analyses.

Figure 5-6 shows that ln(NH_3) increases with increasing temperature. This pattern is consistent across animal types and sites, except for the points at the bottom of the graphs that form horizontal lines where ln(NH_3) is approximately -3. These points correspond to very low NH_3 emissions (near zero), and the fact that they do not follow the same pattern with respect to the relationship to *ta* brings their validity into question.

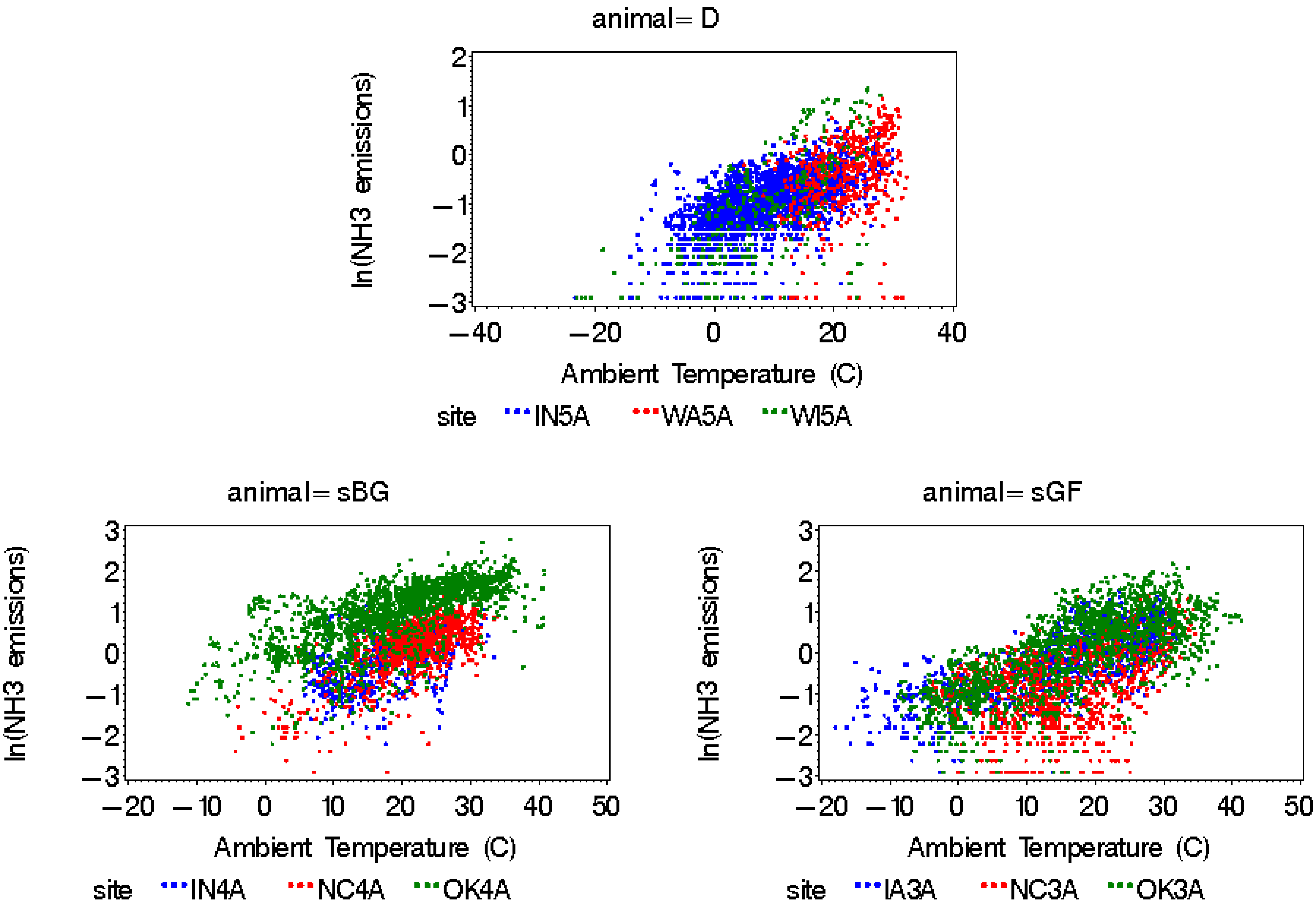


Figure 5-6. NH_3 Emissions vs. Temperature, by Animal Type (color-coded by site)

The axes in three plots of Figure 5-6 were scaled to fit the data for the different animal types. This scaling makes it clear that the range of NH_3 emissions is quite different for the different animal types. It also highlights differences in the range of temperature values covered at the three different types of farms. There are very few temperature values beyond 33 ^{o}C for the dairy sites, while there are quite a few at the swine sites.

To examine the need for an interaction between *ta* and the other meteorological variables, the EPA produced graphs of ln(NH_3) versus *ta* disaggregated by *wsbin* and *habin*. Figure 5-7 shows plots of NH_3 emissions versus *ta* for each of the eight *ws* bins given in Table 5-7. Each bin plot shows a linear regression line with a light blue band indicating the confidence limits for the mean of ln(NH_3) for the given temperature, and dark blue lines for the prediction limits for an individual prediction of ln(NH_3) for the given temperature. These plots allow examination of the change in the slope as *ws* increases. In bins 1 through 5, which are the only bins with sufficient data for trends to be examined, the slope of ln(NH_3) versus *ta* increases with increasing *ws*. Thus, the EPA expected an interaction term formed by multiplying *ws* and some function of *ta* to be helpful in describing variability in NH_3 measured emissions.

Figure 5-8 shows plots of ln(NH_3) versus *ta* for each of the eight *ha* bins given in Table 5-7. Ignoring bin 1, which does not contain sufficient data for drawing conclusions, the slope of ln(NH_3) versus *ta* decreases as *ha* increases. This observation is consistent with the expected result that as *ha* increases, NH_3 emissions from the lagoon decrease.

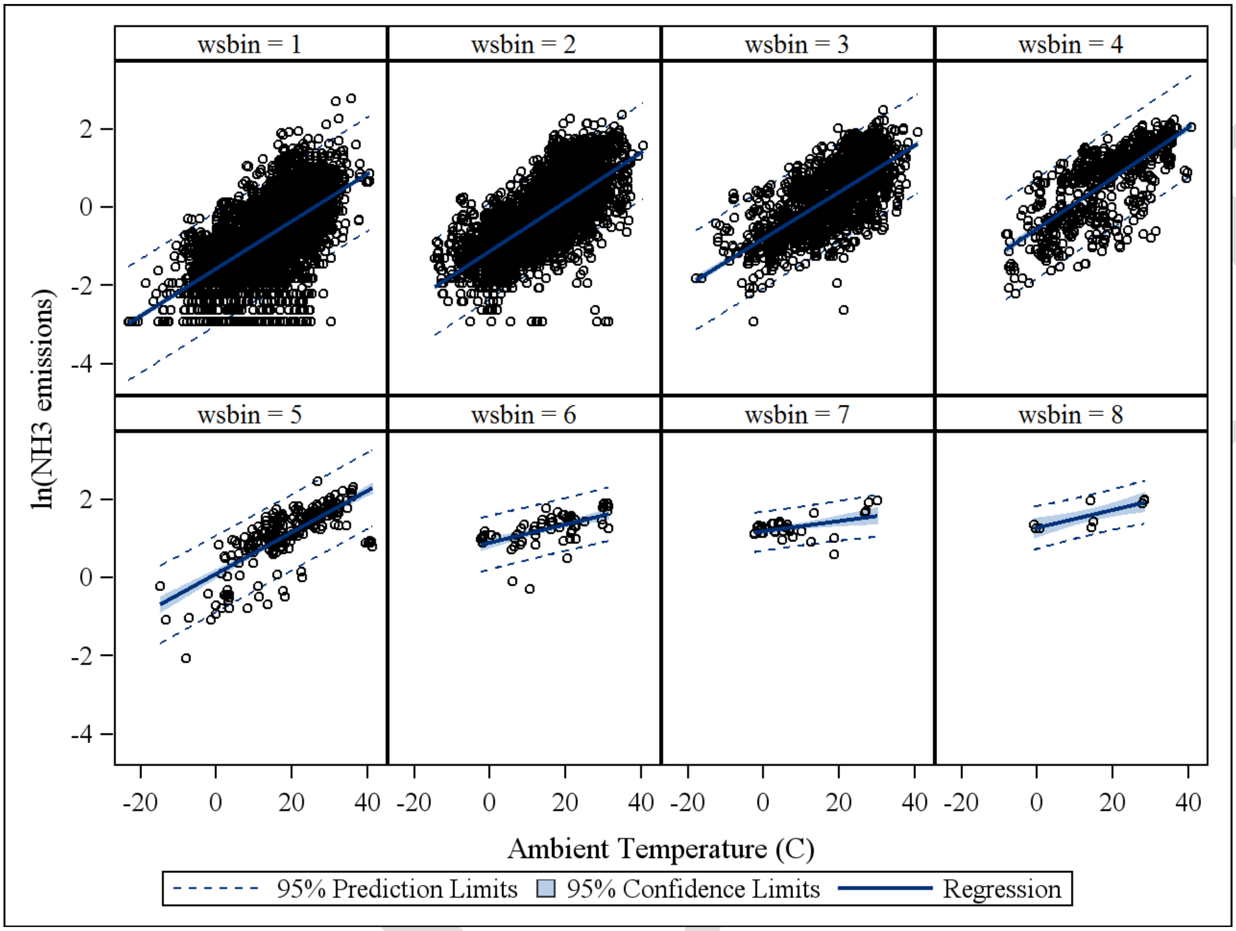


Figure 5-7. Scatter Plots of NH_3 Emissions vs. Temperature, by Wind Speed Bin (*wsbin*)

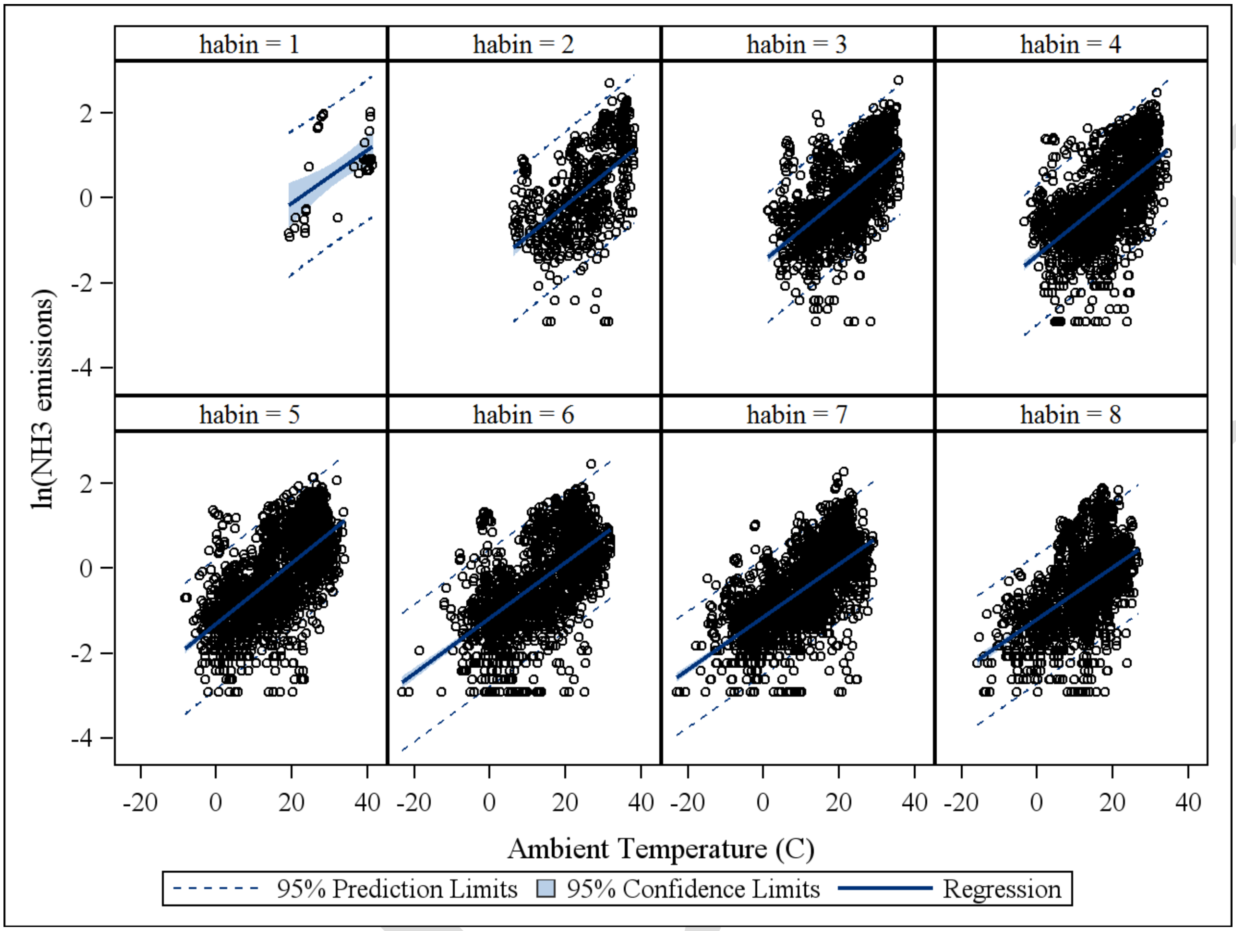


Figure 5-8. Scatter Plots of NH_3 Emissions vs. Temperature, by Relative Humidity Bin

5.3.2 Relative humidity

Figure 5-9 shows that the relationship between ln(NH_3) and *ha* for each of the three animal types is linear. This figure shows that ln(NH_3) decrease slightly with increasing relative humidity. This pattern is consistent across animal types and sites, except for the points at the bottom of the graphs that form horizontal lines where ln(NH_3) is approximately -3. These points correspond to very low NH_3 emissions (near zero), and the fact that they do not follow the same pattern with respect to the relationship to *ha* brings their validity into question.

Coverage of humidity is different for dairy and swine sites. The dairy sites have few values below 15 percent and none below 10 percent, while the swine sites have several values in both ranges.

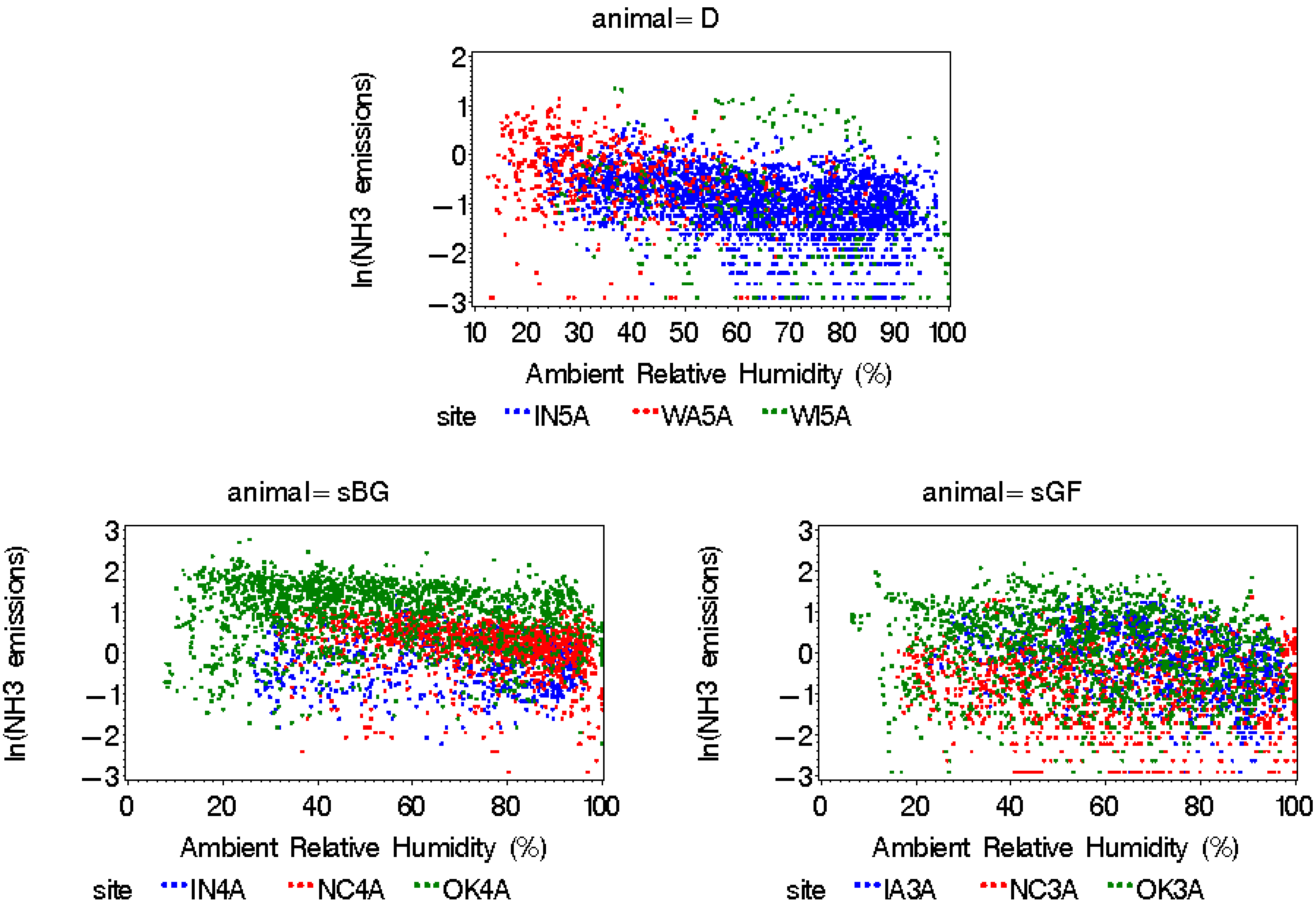


Figure 5-9. NH_3 Emissions vs. Relative Humidity, by Animal Type (color-coded by site)

The EPA investigated the need to include an interaction between NH_3 and *ha* in the mean trend function using scatter plots disaggregated by *ta* and *ws* bins. Figure 5-10 shows ln(NH_3) versus *ha* for the *ta* bins given in Table 5-7. Bins 3 through 7 have sufficient data to examine trends. In bins 3 to 6, the slope describing the relationship between ln(NH_3) and *ha* progresses from a negative slope, to a slope approximating zero, and to a positive slope. The relationship between the slope of *ha* and *ta* from bins 3 through 6 is consistent with the expectation that emissions increase with increasing values of temperature. Based on this analysis, the EPA expected that an interaction between *ha* and *ta* would help explain variability in measured NH_3 emissions. The negative slope shown in bin 7 may be due to a lack of data for the higher values of *ha* in that bin, combined with the effect of the low outliers at the higher values of *ha*.

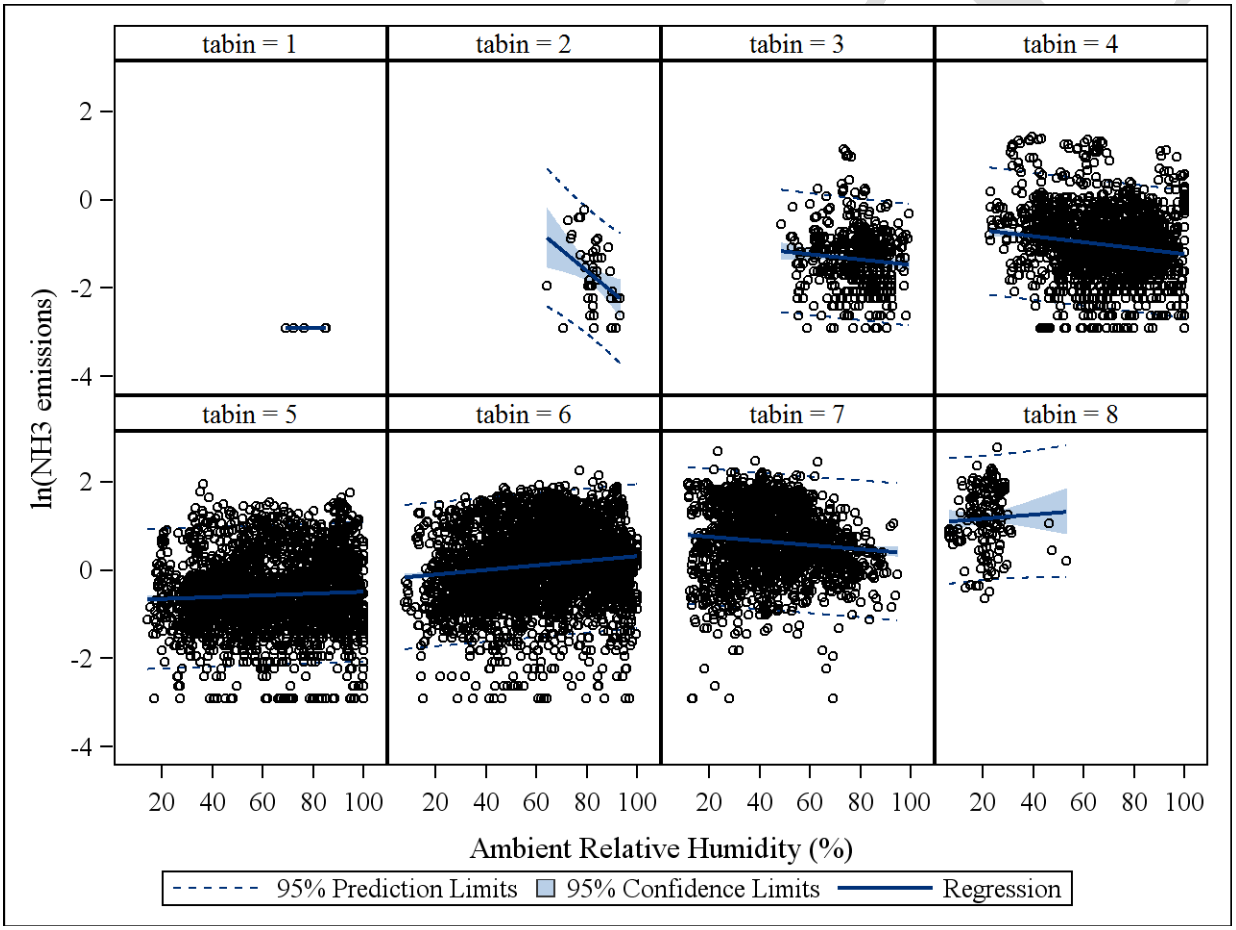


Figure 5-10. NH_3 Emissions vs. Relative Humidity, by Temperature Bin

Figure 5-11 shows ln(NH_3) versus *ha* for each of the *ws* bins given in Table 5-7. In all plots, the effect of *ha* on ln(NH_3) is negative. Bins 1 through 5 have sufficient data to examine trends in *ha*. As the plots progress from lower to higher wind speeds, the steepness of the slope increases (i.e., the slope becomes more negative with increasing *ws*). Notice that there are fewer values of high humidity when *ws* takes higher values, and the suspect near-zero NH_3 emissions that form the horizontal lines at the bottoms of some plots do not occur at the higher values of *ws*. Thus, the EPA expected an interaction between *ha* and *ws* to help explain variability in emissions. Based on this analysis, the candidate mean trend variable that the EPA used for the effect of *ha* on NH_3 emissions is simply a linear function of *ha*. The EPA expected the effect of *ha* to be negative when the values of all other variables are held equal.

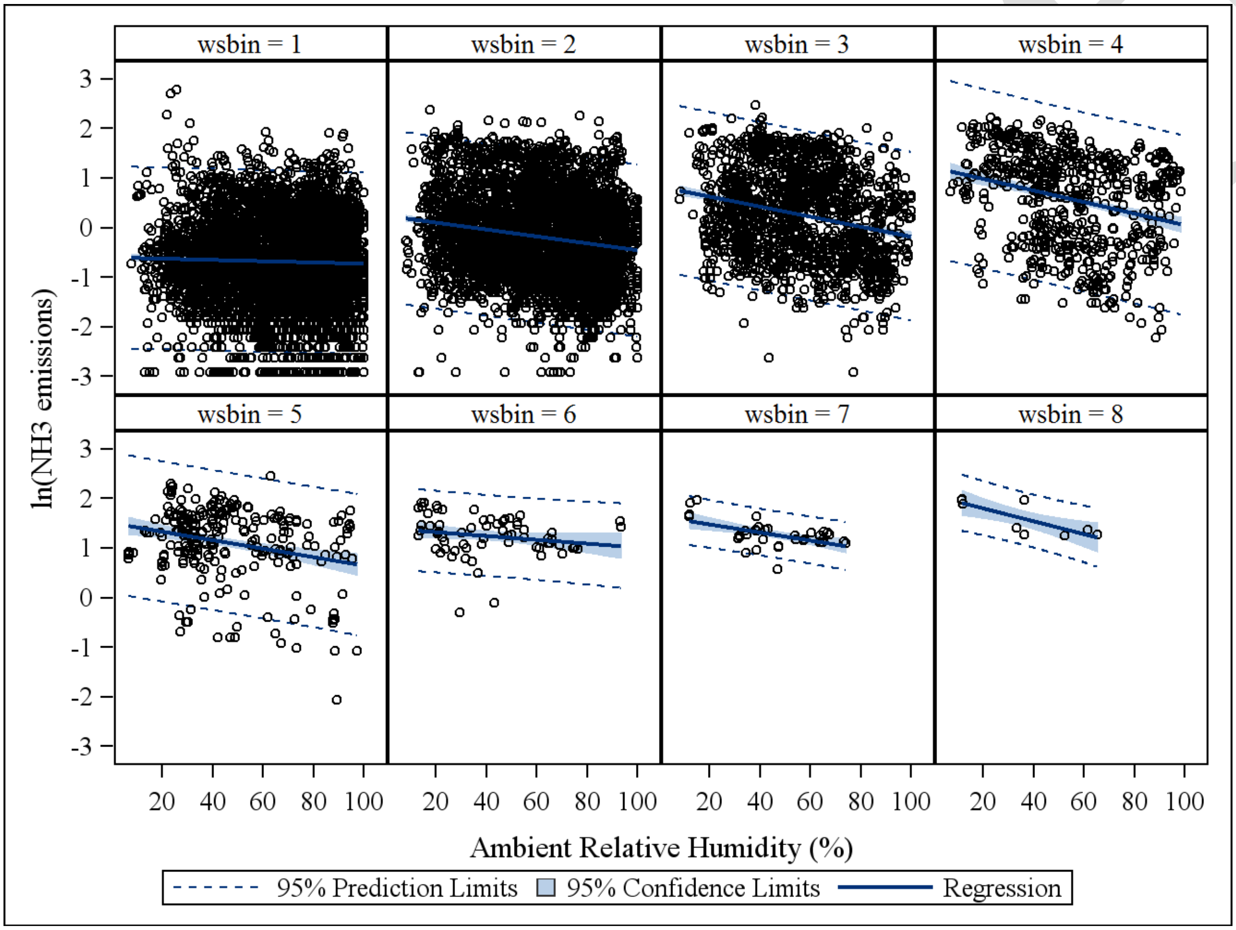


Figure 5-11. NH_3 Emissions vs. Relative Humidity, by Wind Speed Bin

5.3.3 Wind speed

In Figure 5-12, the plots of $\ln(NH_3)$ versus wind speed for each animal type and site show that at all sites, $\ln(NH_3)$ increases with increasing wind speed via an upward sloping curve that is quite steep at low winds speeds, but flattens out as wind speed increases. Notice that the curves are shifted to higher values of $\ln(NH_3)$ for some sites and to lower values for others. Because the meteorological conditions covered by sites WA5A and WI5A are quite different, with WA5A covering more high-temperature conditions and WI5A more low-temperature conditions, the first subplot suggests that the shifts in the curves result from different conditions with respect to other predictor variables.

The four high outliers at low wind speeds for site OK4A occur on days where temperature is very high and humidity very low. Thus, while high wind speed may increase emissions, it is not a necessary condition for high emissions. The expectation that NH_3 emissions would increase as a function of *ws* up to a threshold value, at which they will drop off dramatically is not supported in general, but when these plots are further disaggregated by site, this phenomenon does appear in the scatter plot (not shown) for swine site IA3A. For all observations at site IA3A for which *ws* is high but emissions are low, however, *ta* is less than 20 degrees C, and for many of those observations, *ha* is high. Thus, even for high values of *ws*, high humidity and low temperature can suppress emissions of NH_3.

The differences in the wind speed ranges for dairy sites versus swine sites are even more dramatic than the differences in the ranges of temperature and humidity. The dairy sites do not have wind speed values greater than 10 m/s, while the swine sites have many, especially the breeding and gestation facilities. Because the relationship between NH_3 emissions and wind speed is strong, these differences may have important implications for the ability of the EEMs to learn about the effects of high wind speeds for dairy sites.

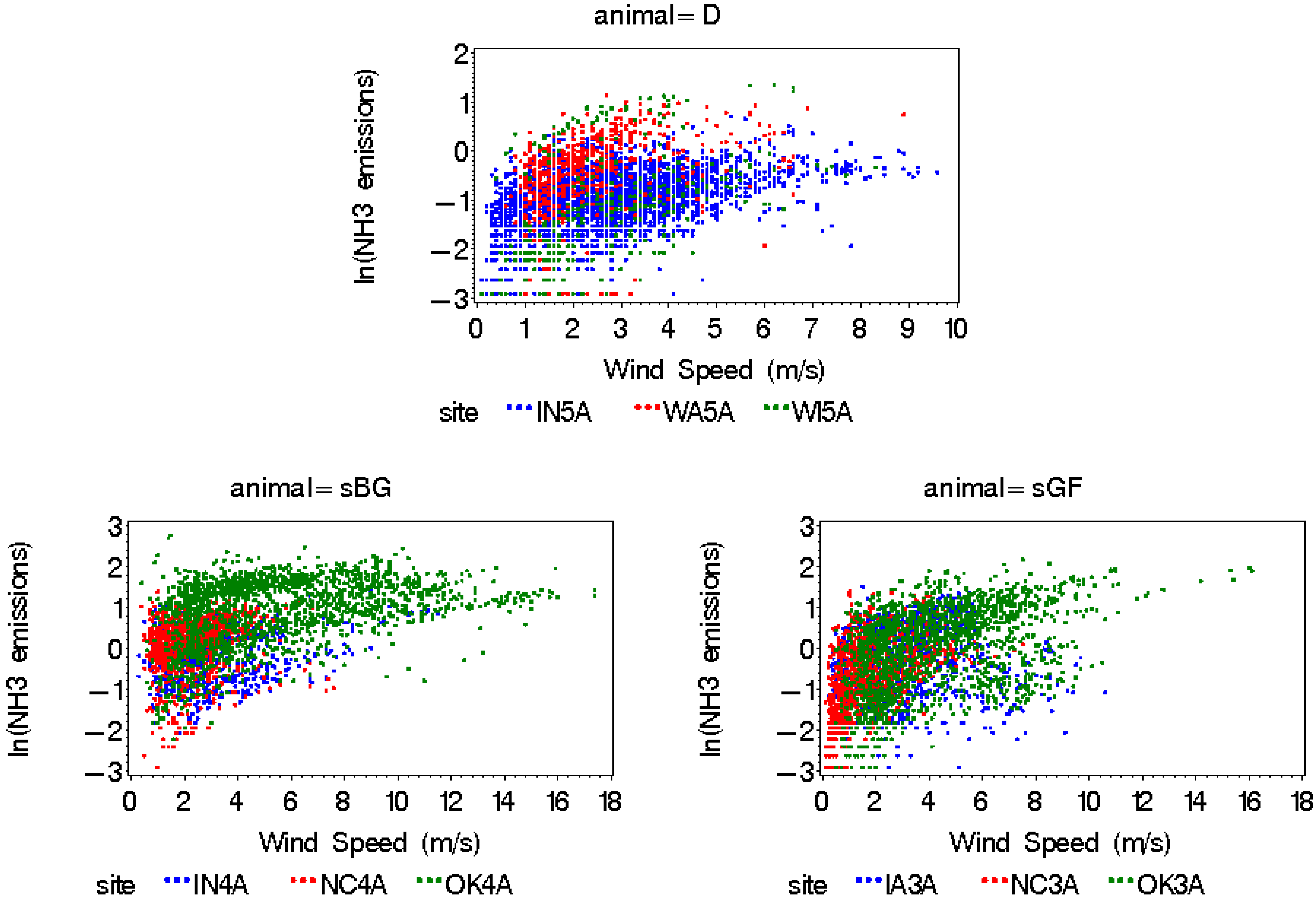


Figure 5-12. NH_3 Emissions vs. Wind Speed, by Animal Type (color-coded by site)

Figure 5-13 shows ln(NH_3) versus *ws* for the *ta* bins given in Table 5-7. Bins 3 through 8 provide sufficient data to determine trends in *ws*. Although these plots do not show changes in the slope describing NH_3 emissions versus wind speed for increasing ranges of *ta*, the EPA still considered an interaction between *ws* and *ta*.

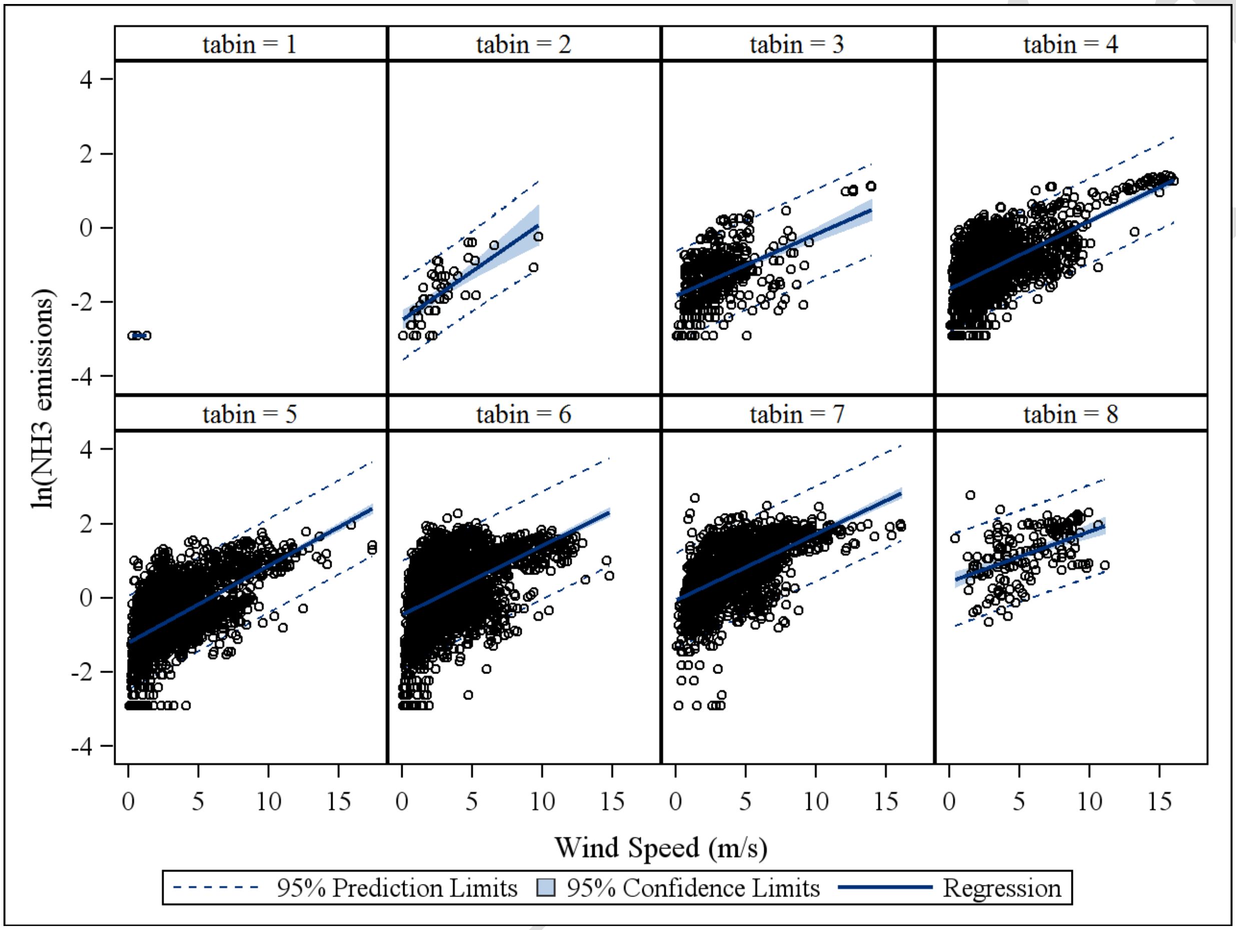


Figure 5-13. NH_3 Emissions vs. Wind Speed, by Temperature Bin

Figure 5-14 shows ln(NH_3) versus *ws* for the *ha* bins given in Table 5-7. The fact that the slope of the regression line decreases slightly with each successive *ha* bin is consistent with the expectation that higher humidity levels block emissions. Despite the curvature in the relationship between ln(NH_3) and *ws*, when the EPA used an exponential function of *ws*, the regressions did not fit as well as when a linear function was used. Therefore, the candidate mean trend variable the EPA used for the effect of wind speed was a linear function of *ws*.

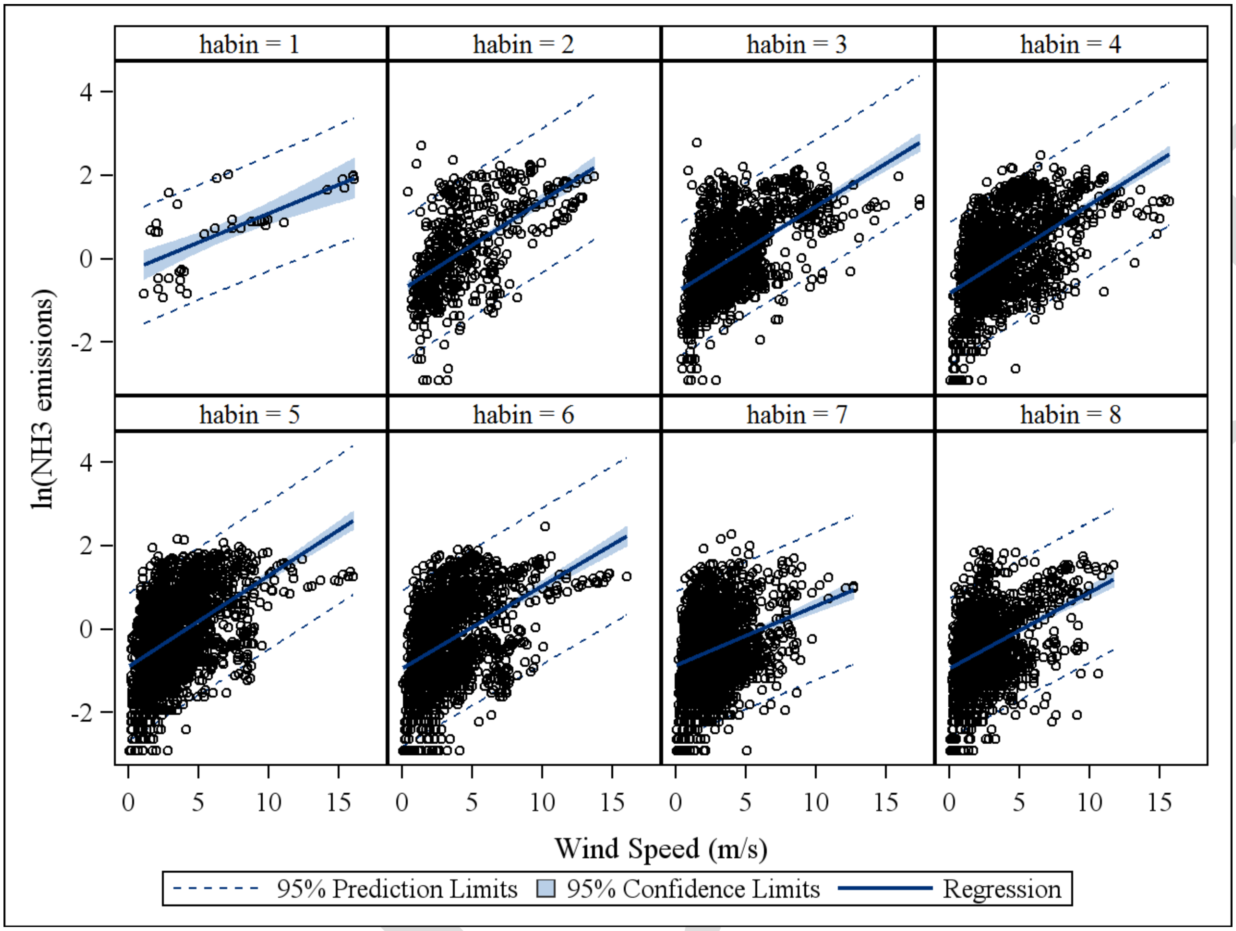


Figure 5-14. NH_3 Emissions vs. Wind Speed, by Relative Humidity Bin

5.3.4 Solar radiation via the diurnal cycle

When the EPA produced exploratory plots to determine what function of *hour* should be used to capture the effect of the diurnal solar-radiation cycle, some data issues became apparent. The box plots in Figure 5-15 show the distributions of ln(NH_3) for each hour of the day from hour 0 (midnight) to hour 23 (11 pm). For all sites, the time of day was reported in local standard time (LST) to mitigate any effect of one state observing daylight savings while another did not. This analysis showed that the highest of the ln(NH_3) maxima occur during the low- or no-sunlight hours of 1, 2 and 21 (1 a.m., 2 a.m. and 9 p.m.), and the lowest occurs at 11 am. This pattern was not consistent with the EPA's expectation that the maximum lagoon NH_3 emissions would occur during daytime hours.

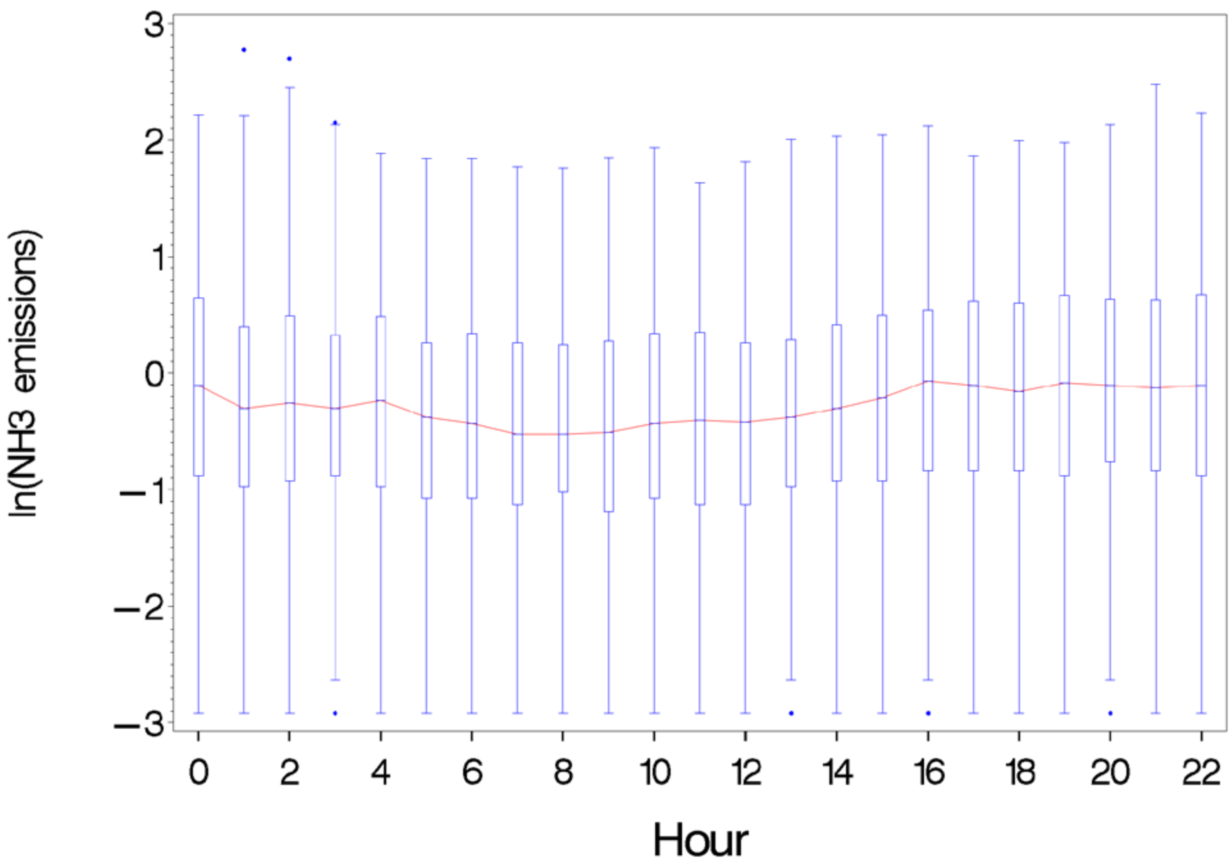


Figure 5-15. Box Plots of NH_3 Emissions vs. Hour

To investigate diurnal patterns more closely, and to keep the effects of extreme outliers from driving conclusions regarding trends, the EPA produced plots of other summary statistics for each hour: the mean, first quartile, median, third quartile, and 90th and 95th percentiles of ln(NH_3) (see Figure 5-16). All of these statistics demonstrate a similar periodic diurnal cycle for which the curvature appears sinusoidal. The hours with the lowest values for each statistic fall between 7 and 13 (7 am to 1 pm), and the hours with the highest fall between 20 and 0 (8 pm and midnight). This analysis was also not consistent with the expectation that effect of *hour* on NH_3 emissions would be due to the effect of solar radiation on lagoon temperature. Although the peaking of emissions at night might be due to a lagged effect of solar radiation on lagoon temperature, the EPA plotted other variables as a function of *hour* to investigate.

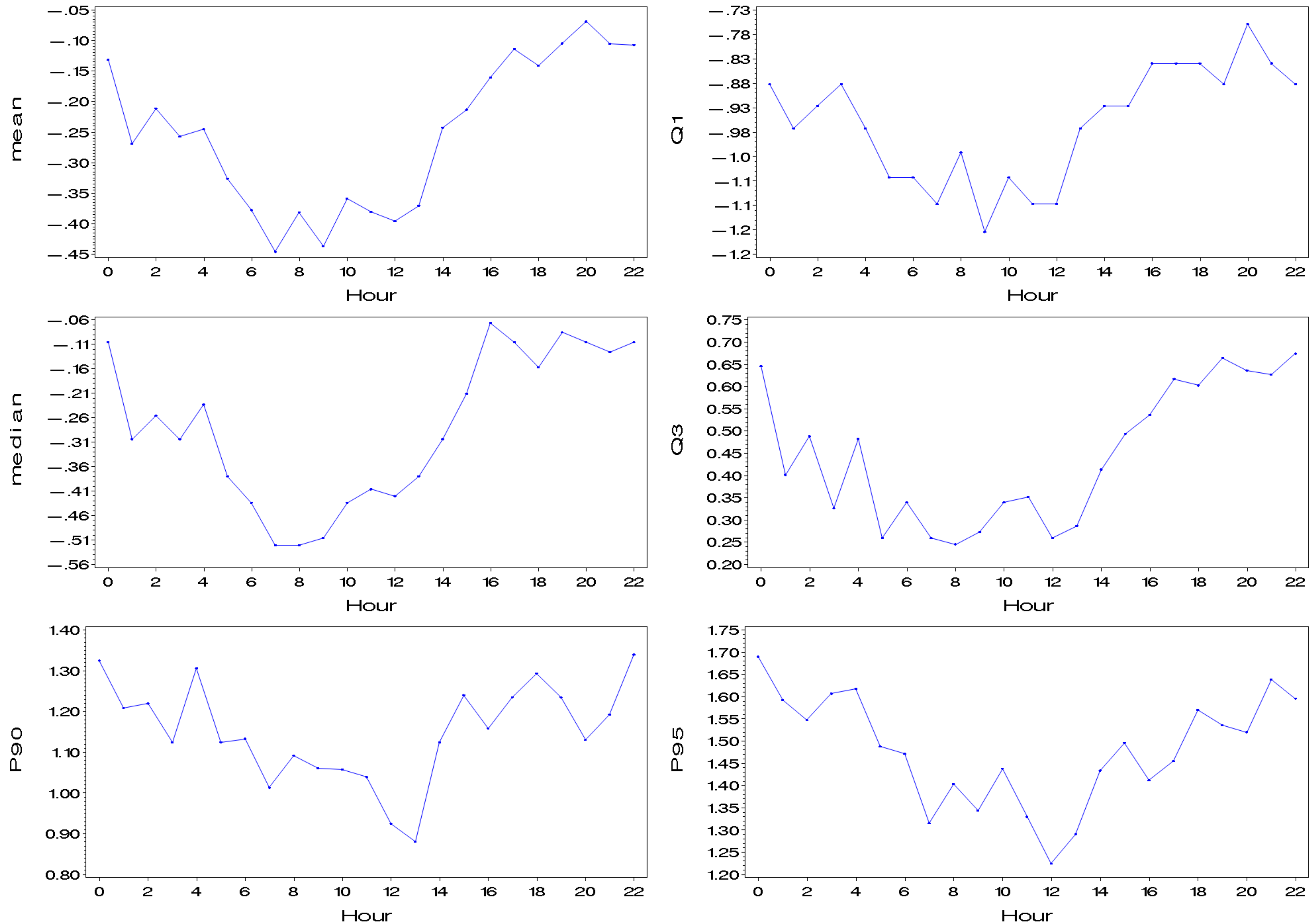


Figure 5-16. Summary Statistics for NH_3 Emissions vs. Hour

Figure 5-17 and Figure 5-18 show the distributions of *ta* and *ha* by *hour*, respectively. These plots show that the peak temperature values occurred over night and the peak humidity values occurred in the middle of the day. These patterns are counter intuitive and might indicate that there is an issue related to the time values in the data submitted to the EPA. The EPA continued using the values of *hour* as reported in the data submitted to the EPA. The EPA plans on resolving this issue with the NAEMS researchers. Correcting the values of *hour*, if needed, will be a minor change, and if the values of *hour* are shifted by 12, that will not affect the selection of functional form through which *hour* should enter the mean trend function, but it will affect the final regression coefficient estimates. Thus, if the values of *hour* change, the EPA will simply refit the EEM to the revised data to obtain the appropriate regression coefficient estimates.

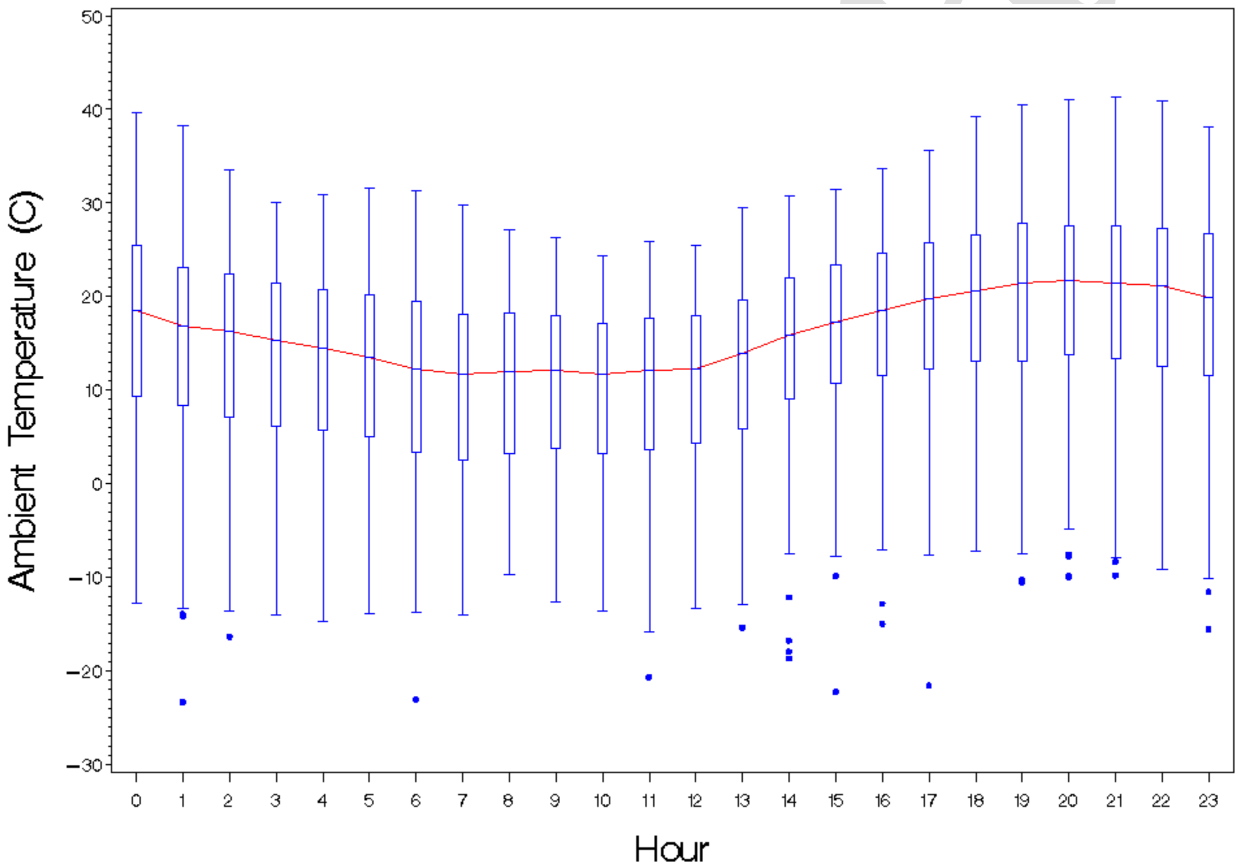


Figure 5-17. Box Plots of Temperature vs. Hour

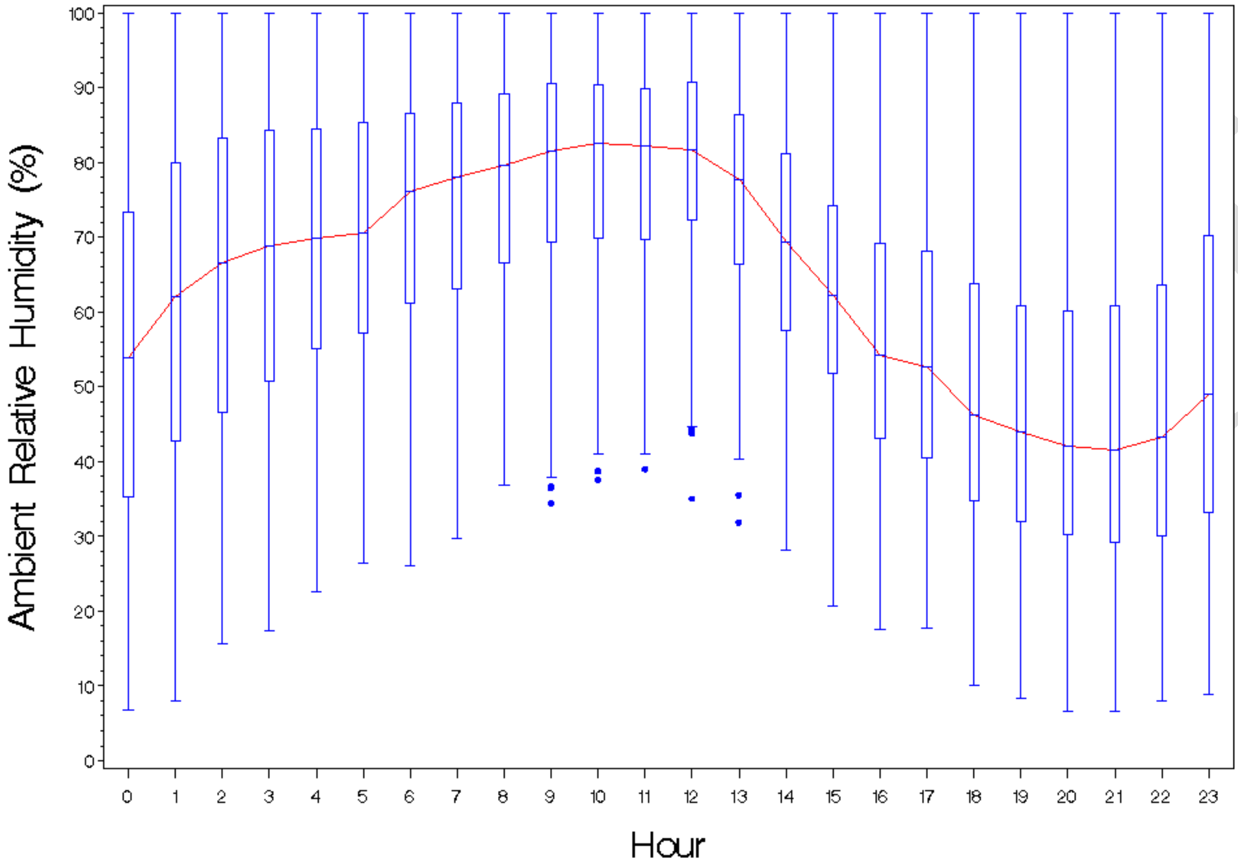


Figure 5-18. Box Plots of Humidity vs. Hour

Because representing a sinusoidal function of *hour* with a 24-hour period requires a sine and a cosine term, the EPA created the two new variables, *sh* and *ch*, where:

$$sh = \sin\left(\frac{2\pi\,(hour)}{24}\right) \quad \text{and} \quad ch = \cos\left(\frac{2\pi\,(hour)}{24}\right) \qquad \text{Equation 5-2}$$

The EPA used *sh* and *ch* as the variable names to indicate that these are hourly sine and cosine terms, as opposed to the daily sine and cosine terms *sd* and *cd* used to model an annual cycle, as explained in the next section. Thus, the candidate mean trend variables the EPA used for the effect of a diurnal solar radiation effect on NH_3 emissions were *sh* and *ch*.

5.3.5 Solar radiation via the annual cycle

To capture the effect of the annual solar radiation cycle on NH_3 emissions, the EPA used the variable *jday*, defined as the Julian day, which takes values 1 through 366 in leap years, and 1 through 365 otherwise. The plot of NH_3 emissions versus *jday* in Figure 5-19 at first makes it appear that not only is there an annual cycle; there is also a semi-annual cycle. There is a peak in

NH_3 emissions for the days corresponding to summer months and a minimum for days corresponding to winter months, but there are also smaller peaks for days corresponding to spring and fall months.

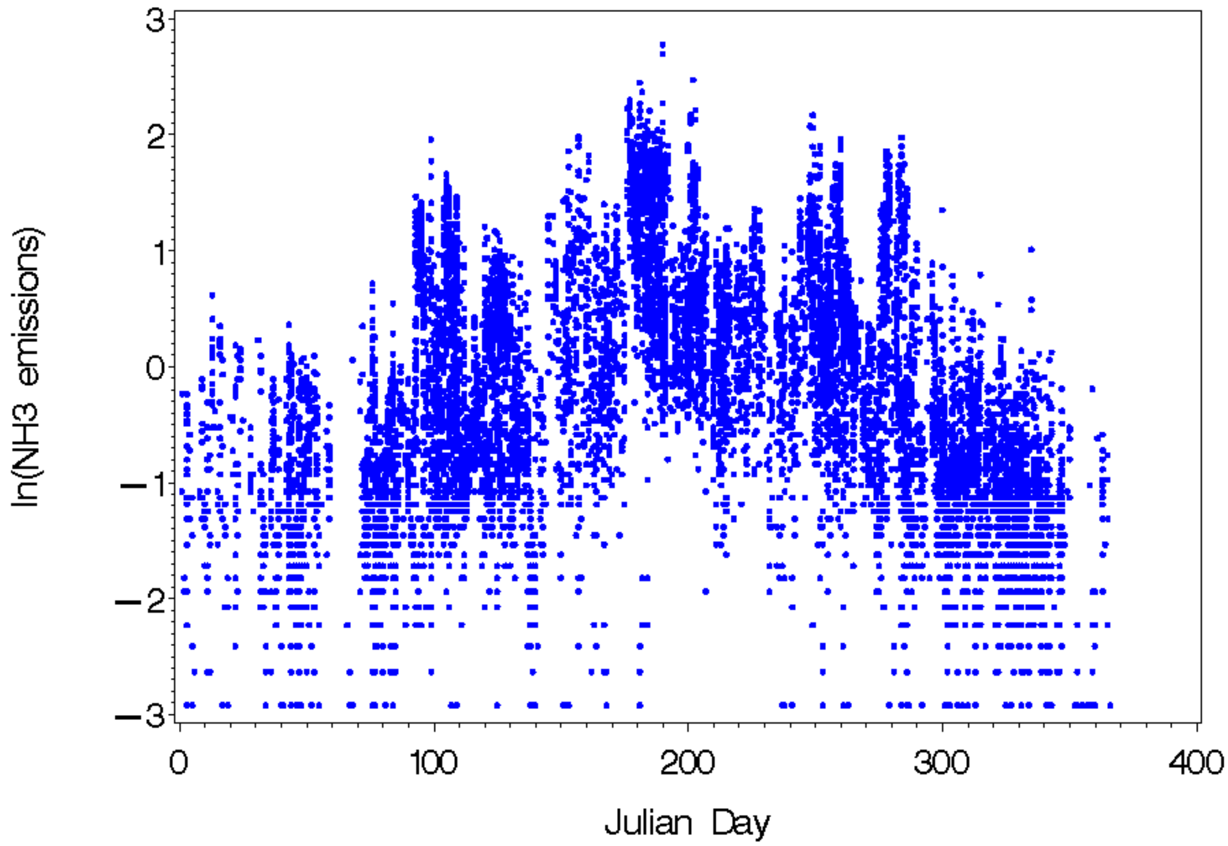


Figure 5-19. NH_3 Emissions vs. Julian Day

A histogram of the value of the variable *month* (Figure 5-20) shows that one explanation for the secondary peaks is that the majority of NH_3 emissions observations occur in the spring and fall. Because there are more observations during these seasons, there are more opportunities to see extreme values.

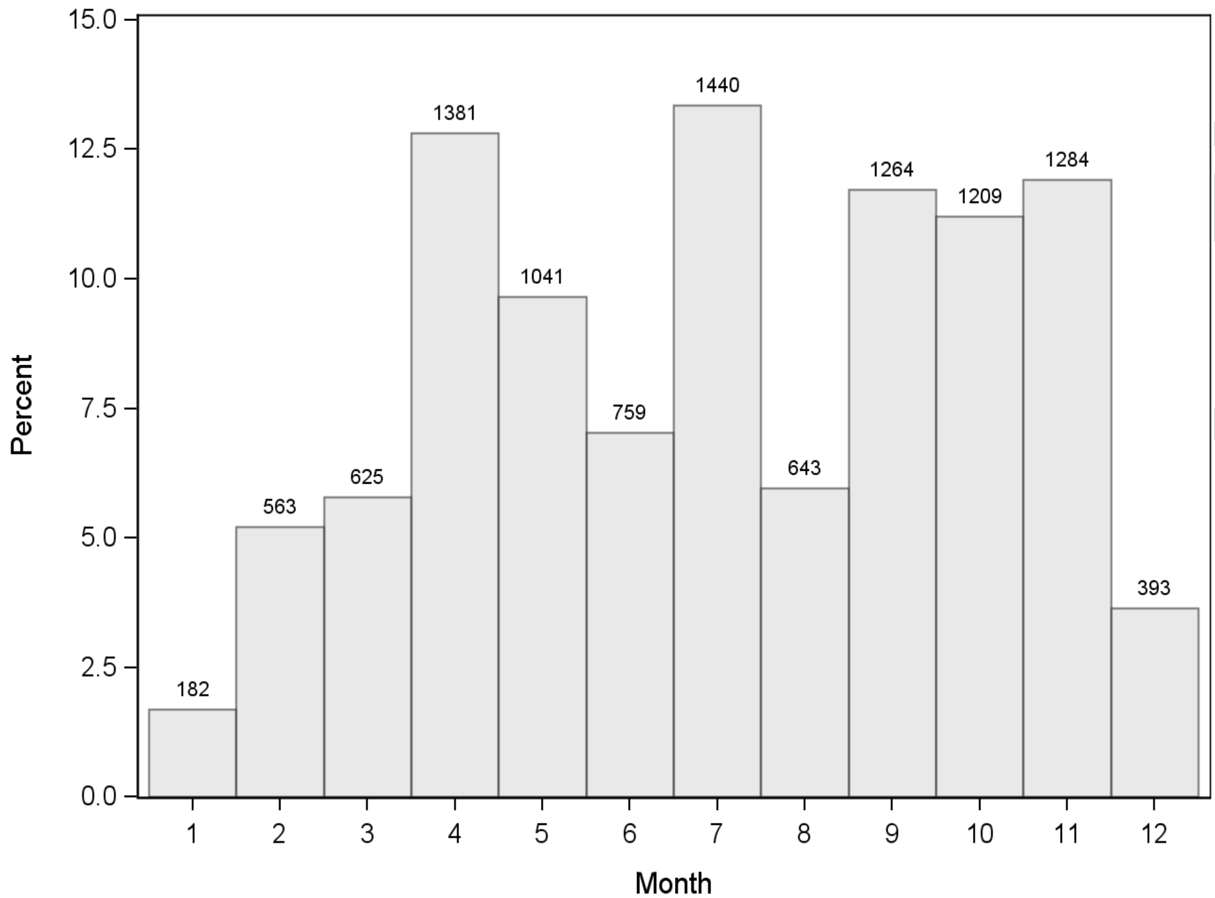


Figure 5-20. Percent and Number of Observations per Month

Based on the patterns shown in Figure 5-19 and Figure 5-20, the EPA chose to represent the annual effect of solar radiation on NH_3 emissions as a sinusoidal periodic function of *jday* with a period of 365.25 days. Therefore, the EPA created the two new variables, *sd* and *cd*, where:

$$sd = \sin\left(\frac{2\pi\,(jday)}{365.25}\right) \quad \text{and} \quad cd = \cos\left(\frac{2\pi\,(jday)}{365.25}\right) \qquad \text{Equation 5-3}$$

To ensure that additional harmonics were not needed, the EPA also created variables to represent a semi-annual cycle of solar radiation. These variables are the same as those in Equation 5-2, except that the denominator is 182.625. An ordinary least squares regression of NH_3 emissions on the first two terms (i.e., *sd* and *cd* on an annual basis) produced an R^2 of 0.2334. A regression on all four terms (i.e., *sd* and *cd* on annual and semi-annual bases) produced an R^2 of 0.2450. Because the semi-annual cycle explained less than two percent of variability beyond what was explained by the annual cycle, the EPA did not include the two semi-annual cycle terms as mean trend variables. Thus, the candidate mean trend variables used for the effect of the annual solar radiation cycle NH_3 emissions were *sd* and *cd*.

5.3.6 Farm-Based Variables

In Section 5.1, the EPA's initial decisions regarding what farm-based variables to include as candidate predictor variables were made based on availability and the potential for the variable to influence NH_3 emissions. This section discusses the decisions regarding variable selection that the EPA made based on the ability of the EEMs to discern the effects of a variable on NH_3 emissions. This section also presents two challenges that were encountered in this analysis.

Table 5-10 shows the farm-based predictor variables values that describe the farm and lagoon configuration. Each of the variables is designated as being a static categorical or continuous/discrete predictor variable. The variables labeled continuous/discrete were treated as continuous variables in developing the EEMs, but were available as discrete numeric variables obtained from the NAEMS supporting documentation (e.g., SMPs). The continuous/discrete variables can take, at most, nine different values (i.e., one value for each site). However, because the EEMs must apply to lagoons other than those that participated in the NAEMS, a continuous function of these variables must be used in the EEMs.

Table 5-10. Summary of Farm-Based Predictor Variables

Variable [a]	Variable Type	Dairy Sites			Swine Sites					
		IN5A	**WA5A**	**WI5A**	**IA3A**	**NC3A**	**OK3A**	**IN4A**	**NC4A**	**OK4A**
animal[b]	Categorical	1	1	1	-1	-1	-1	-1	-1	-1
mms	Categorical	S	FS	FS	PP	PPR	PPR	PP	PPR	PPR
odorctrl	Categorical	N	N	N	N	N	Y	N	N	N
inlets	Categorical	1	1	1	2	1	1	1	1	2
*capacity**	Cont/Disc	2,600	5,600	1,700	3,840	8,000	3,024	1,400	2,000	2,784
*adultwt**	Cont/Disc	1,500	1,400	1,400	150	135	170	475	433	490
*barns**	Cont/Disc	4	6	6	4	5	3	9	9	3
*sa**	Cont/Disc	106	141	77	25	204	121	121	250	242

[a] For the variable *mms*: S = scrape, = FS = free-stall flush, PP = pull-plug pit, PPR = pull-plug w/pit recharge. An asterisk (*) indicates that the variable has not yet been centered and scaled.
[b] For the variable *animal*: 1 is used for dairy sites and -1 is used for swine sites.

5.3.6.1 ***Challenges due to small number of NAEMS sites***

The first challenge is that there are only nine sites, or nine data points, but there are eight farm-based predictor variables. Although there are many observations of the continuous variables over time at each of the nine sites, the value of the farm-based predictor variables is the same for a given site over all time points. Effectively, this means that there are only nine data points available from which to learn about the effects of farm-based predictor variables on NH_3 emissions. It is impossible to learn about the effect of eight variables on NH_3 emissions with only nine data points. The following example demonstrates the challenge.

The example assumes that the predictor variable of interest is surface area, *sa*, and that only one half-hour period is being considered. For the example, the assumptions were made that the meteorological conditions are exactly the same at all sites and that data are available for only two sites. The surface area at Site 1 is 121,000 ft^2 and 200,000 ft^2 at Site 2. The NH_3 emissions from Site 1 are 1.0 kg and 1.5 kg from Site 2.

Equation 5-1 can be modified for this simplified example by omitting the subscript t, not including the optional term e_s and using the identity link function. The simplified equation is given in Equation 5-4.

$$\boldsymbol{Y_s \sim Gamma\,\{\mu_s, \phi\};\;\; \mu_s = \beta_0 + \beta_1 sa_s; s = 1, 2} \qquad \text{Equation 5-4}$$

Using the example values for surface area and NH_3 emissions, the simplified equation (i.e., EEM) can be written as Equation 5-5 and Equation 5-6, respectively.

$$\boldsymbol{1.0 \sim Gamma\,\{\mu_s, \phi\};\;\; \mu_s = \beta_0 + \beta_1 121} \qquad \text{Equation 5-5}$$

$$\boldsymbol{1.5 \sim Gamma\,\{\mu_s, \phi\};\;\; \mu_s = \beta_0 + \beta_1 200} \qquad \text{Equation 5-6}$$

As illustrated in these site-specific equations, there are three unknown parameters to be estimated from this dataset of two observations: β_0, β_1, and ϕ. Three unknowns cannot be solved using only two equations. In this example, the number of parameters to be estimated (i.e., the model degrees of freedom), exceeds the number of observations in the dataset (i.e., total degrees of freedom). Consequently, it is impossible to estimate the parameters in the EEM.

The purpose of the scaling parameter ϕ can be thought of as quantifying variability around the mean. It determines how wide or narrow a curve like those in Figure 5-5 would be. Ignoring ϕ, solving for β_0 and β_1, and using the resulting EEM would be equivalent to saying that there is a 100 percent probability that all sites have the emissions given by this linear function under the same meteorological conditions. The emissions from different sites, however, will be different even if the meteorology is the same and the values of the variables in the mean trend function are the same because there are different characteristics of each farm that have not been accounted for in the EEM that result in different emissions.

If the example included additional data for Site 3 (surface area = 121,000 ft^2 and NH_3 emissions of 1.1 kg), the three unknowns could be resolved. However, the EEM would perfectly fit these three data points, and thus these three sites. In this instance, although an estimate for the parameter that quantifies variability could be determined, it would not be a good estimate, because there are simply not enough data values from which to learn about variability. If such an EEM were to be applied to sites other than these three (for the same meteorological conditions), the variability quantified by ϕ would likely not be large enough for the 95-percent prediction intervals to contain 95 percent of the true the emissions values.

The difference between the total degrees of freedom and the model degrees of freedom is called the "error degrees of freedom." The error degrees of freedom must be large enough to adequately quantify variability of emissions around the mean. In other words, for any given set of values of predictor variables, the value of emissions will not be the same. Instead, there will be a distribution of values, and the EPA uses the gamma distribution to quantify the range of values of emissions for the chosen predictor variables. This variability is the source of uncertainty when the EEM is applied to sites outside the study. The estimated value of parameters must be based on many observations to quantify the spread of this distribution. The intercept, β_0, uses one degree of freedom, and the scale parameter, ϕ, uses one degree of freedom. One continuous mean trend variable requires one degree of freedom for its slope. The number of degrees of freedom used by each categorical variable is one less than the number of values the variable can take. For example, the categorical variable *mms*, which takes four possible values, uses three degrees of freedom.

With only nine NAEMS sites, there are nine total degrees of freedom with which to learn about farm-based variables. After using two degrees of freedom with the intercept and scale parameter, there are only seven degrees of freedom left. The EPA decided to use two additional degrees of freedom in the mean trend function and allow five error degrees of freedom for quantifying variability. The decision as to how many error degrees of freedom are sufficient is a judgment call, and was assessed by evaluating the percentage of NH_3 emissions in the cross-validation dataset that fall inside the 95-percent prediction interval.

In developing the EEMs from the NAEMS data, if the EPA were to attempt to develop a separate EEM using only data from the dairy sites, there would be only three data points. Consequently, at most, only one farm-based predictor variable could be used. The resulting EEM would perfectly fit the three dairy sites, with no degrees of freedom for error, but it would be inappropriate to use it for predicting other sites due to the inability to adequately quantify uncertainty.

Because these obstacles would restrict individual dairy and swine EEMs to being functions of too few predictor variables, the EPA decided to use data from both dairy and swine lagoons to develop the EEM. The differences in NH_3 emissions from different animals are driven by differences in the contribution of precursors of NH_3 to the lagoon, differences in farm/lagoon management practices and/or differences in the physical and chemical processes that convert precursors to actual emissions. Of all the candidate predictor variables discussed in Section 5.1.1, those describing the lagoon liquid, for which the effects on NH_3 emissions were detailed in Section 5.1.1.2, are the variables whose effects on emissions come closest to representing the quantity of precursors and potential for conversion of precursors to emissions. Unfortunately, as shown in Table 5-3 and Table 5-4, the availability of this information was too restrictive to allow its use in EEM development. Thus, the EPA chose to use the farm-based predictor variables in Table 5-9 as surrogates for those describing differences in the lagoon liquid and thereby accounting for any difference in emissions from dairy and swine lagoons

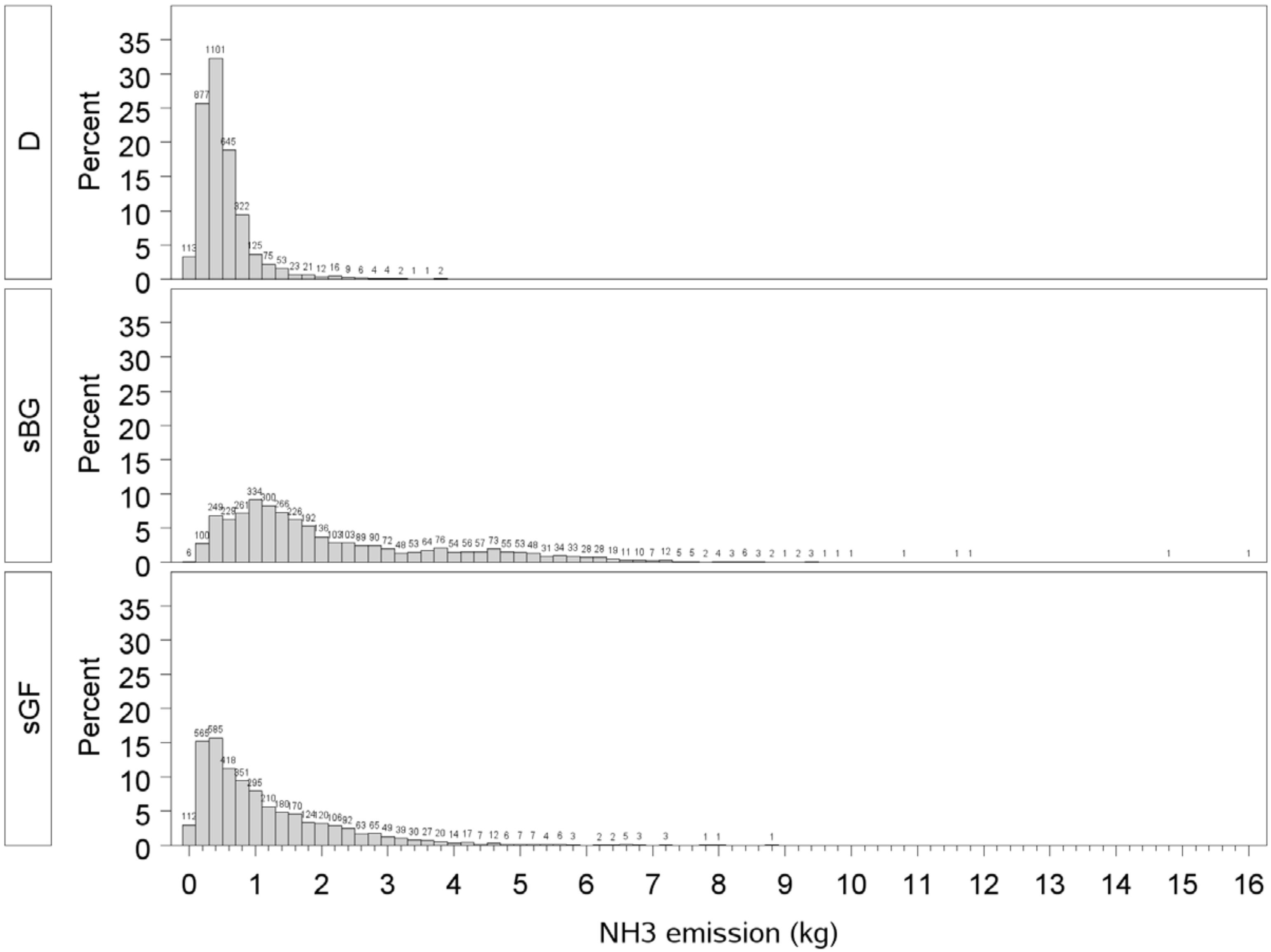
Because the EPA allowed for five error degrees of freedom to quantify variability, only two degrees of freedom remain for the farm-based variables The EPA decided to compare three candidate EEMs, each using two of three variables chosen to represent differences in precursor content and potential for conversion of precursors to emissions. Using the variable *animal* (i.e., swine or dairy) might account for a broad range of differences in the contribution of precursors due to differences in feed, nitrogen excretion and differences in manure management. The variables *capacity* and *adultwt* would also provide information about contribution of precursors to the lagoon. However, because using both of these variables would use up two degrees of freedom, the EPA created the variable *size*, which was defined as the product of *capacity* and *adultwt*. This compound variable quantifies the average live animal mass on the farm while using only one degree of freedom. The NH_3 emissions might also be greater when surface area is larger due to more interaction of the surface of the lagoon with the atmosphere providing more opportunities for desorption of NH_3. Thus, the three variables chosen were *animal*, *size* and *sa;* and the three candidate EEMs each contain two of these three variables.

5.3.6.2 ***Challenges due to unbalanced coverage of meteorological conditions***

In addition to the challenges posed by having only nine data points from which to draw conclusions, the unbalanced representation of meteorological conditions at the different sites may also confound the ability of the EEM to discern effects of farm-based predictor variables. Figure 5-21 shows histograms of NH_3 emissions disaggregated by the three animal types: dairy, swine breeding and gestation, and swine growing and finishing. The number of observations per bar are labeled to make it easier to see the outliers. Notice that the highest value of emissions for any dairy site is around 4 kg, but both swine sites have many values above 4 kg. The highest value of emissions for any swine growing and finishing site is around 9 kg, but the values for the breeding and gestation sites go as high as 16 kg.

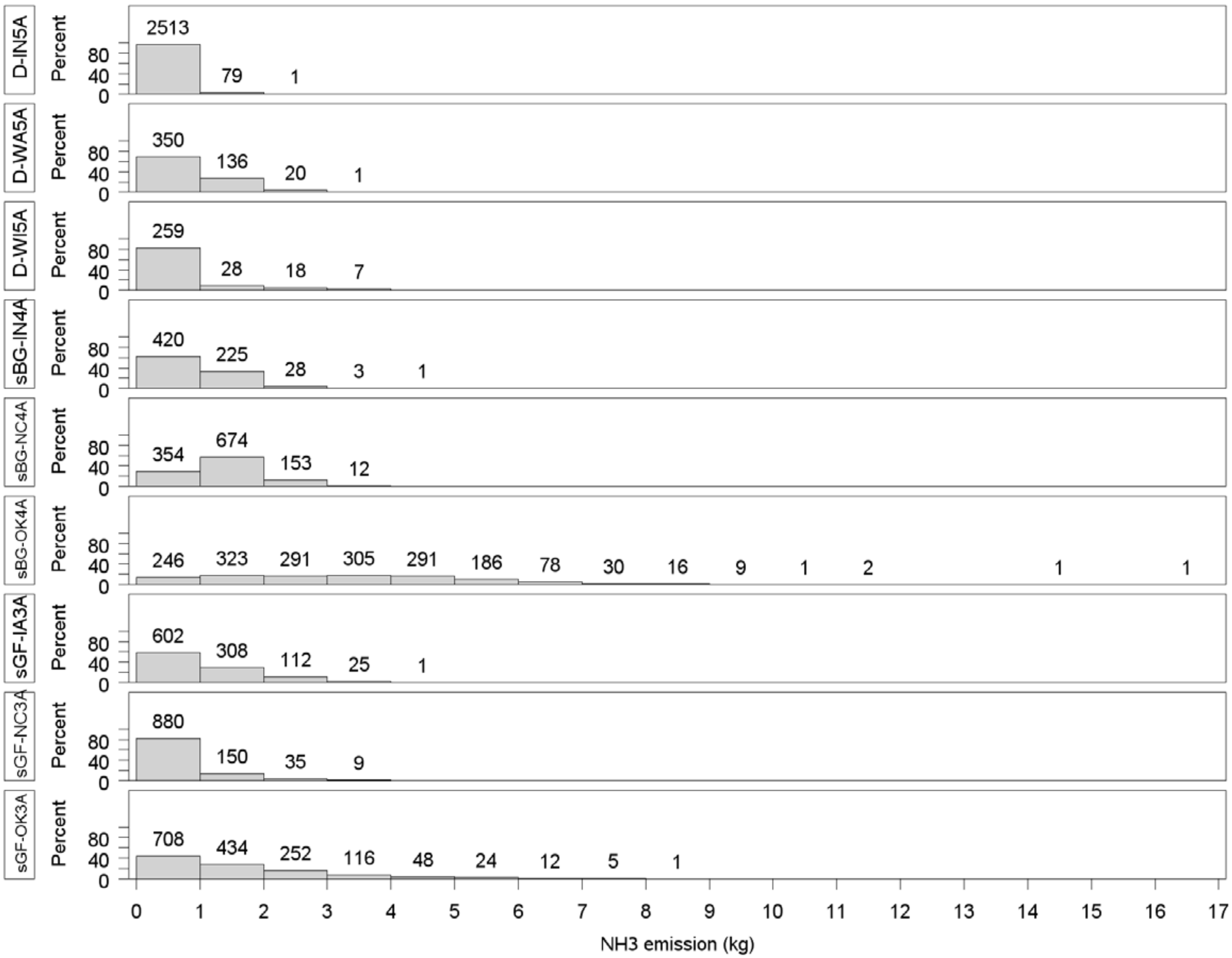
Figure 5-21 gives histograms of NH_3 emissions further disaggregated by both animal type and site. Within each animal type, the mean and variation of emissions tend to be driven by a single site and the differences among sites were due primarily to different seasonal monitoring patterns at each site. Figure 5-22 gives the number of observations per month for each site, with a numerical summary provided in Table 5-11. Figure 5-21 shows that the greatest number of high emissions values among all sites was measured at swine breeding-gestation site OK4A, and Figure 5-22 shows that this site had far more measurements for June and July than any other site. Also, at all three of the breeding-gestation sites, relatively few measurements were made during the colder months, and at IN4A no measurements were taken from November through March. Accordingly, the swine breeding-gestation sites are more representative of lagoon emissions during the warmest six months of the year. In contrast, the great majority of measurements at two of the dairy sites (WI5A and IN5A) were made in the colder months of October through April. At dairy site IN5A, which had the lowest emissions of all sites, no measurements were taken in June, July or August. The highest dairy emissions were at WA5A where almost all of the measurements were taken between June and September.

At this point, no conclusion can be drawn from plots or summary statistics regarding differences in NH_3 emissions from lagoons for different animal-types because the meteorological conditions under which the data were collected for the different sites are so different. The emission differences among animal type appear to be driven by the differences in the NAEMS sampling schedule across the sites, with higher emissions occurring in the summer months and lower emissions in the colder months across all animal types. While the seasonal data availability is different among sites, the NAEMS data collectively provide coverage of all the seasons when all sites are combined, though the data are quite sparse for January.



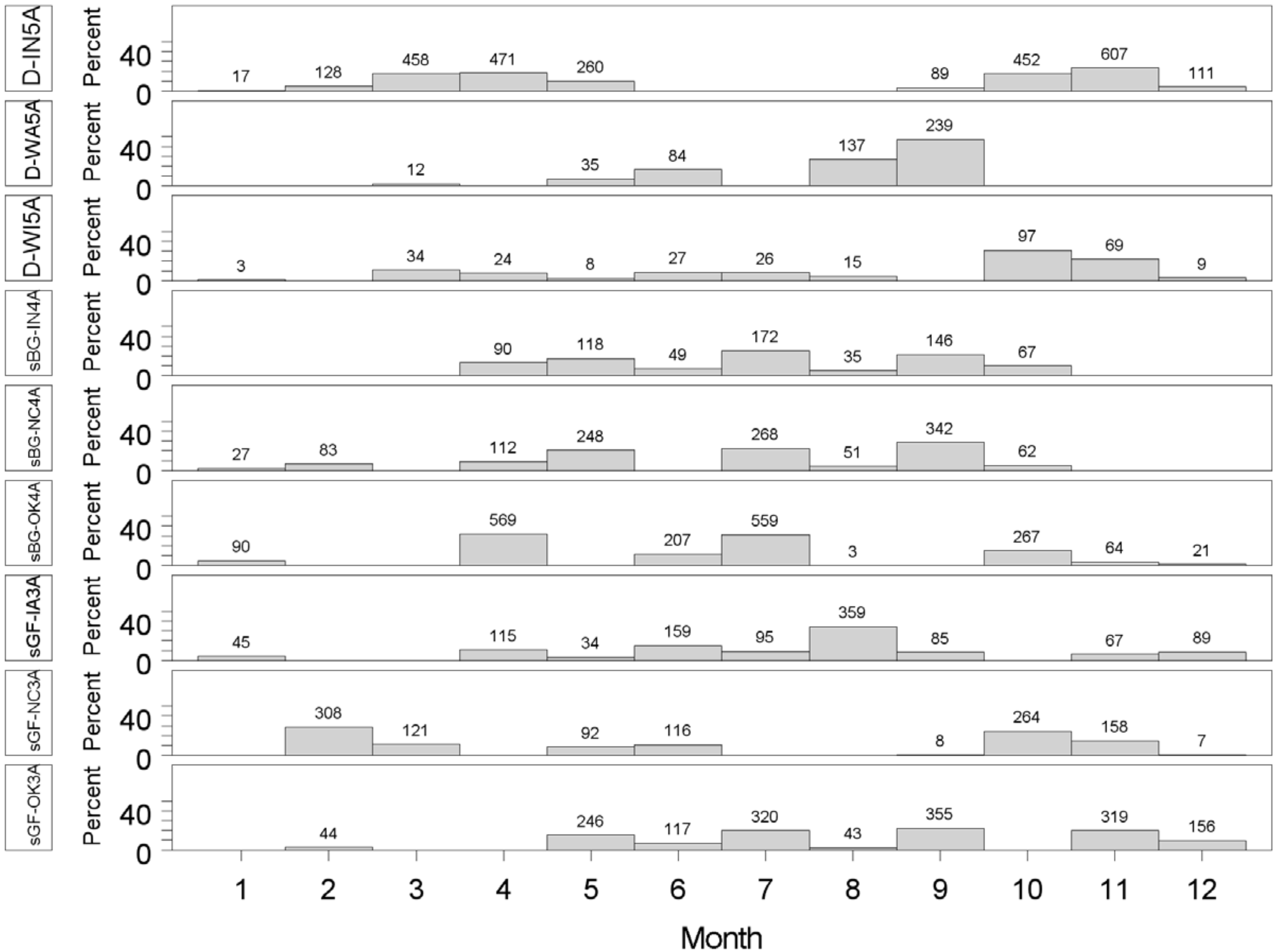
D= Dairy sites; sBG = Swine breeding/gestation sites, sGF = Swine grow/finish sites

Figure 5-21. Histogram of NH_3 Emissions by Animal Type



D= Dairy sites; sBG = Swine breeding/gestation sites, sGF = Swine grow/finish sites

Figure 5-22. Histogram of NH_3 Emissions by Animal Type and Site



D= Dairy sites; sBG = Swine breeding/gestation sites, sGF = Swine grow/finish sites

Figure 5-23. Histogram of Month by Animal Type and Site

Table 5-11. Summary of Data Available by Month and Site

Month	All Sites	IN5A	WA5A	WI5A	IN4A	NC4A	OK4A	IA3A	NC3A	OK3A
1	182	17	-	3	-	27	90	45	-	-
2	563	128	.	.	-	83	-	-	308	44
3	625	458	12	34	-	-	-	-	121	.
4	1381	471	-	24	90	112	569	115	-	-
5	1041	260	35	8	118	248	-	34	92	246
6	759	-	84	27	49	-	207	159	116	117
7	1440	-	-	26	172	268	559	95	-	320
8	643	-	137	15	35	51	3	359	-	43
9	1264	89	239	-	146	342	-	85	8	355
10	1209	452	-	97	67	62	267	-	264	-
11	1284	607	-	69	-	-	64	67	158	319
12	393	111	-	9	-	-	21	89	7	156

Note: Dashes indicate no data was available for the site for the month.

5.3.7 Working mean trend function

To obtain a working mean trend function, the EPA evaluated the relative importance of the meteorological mean trend variables. The EPA fit a model using the gamma distribution and the log link function to: (1) the individual meteorological variables, (2) two pairs of time-based variables representing the main effects of the annual solar radiation cycles, (3) all of these variables simultaneously, and (4) to all two-way interactions among this collection of variables. Table 5-12 presents the fit statistics for each of the four variable combinations. The first row of Table 5-12 gives the estimated slope of the individual meteorological variables. As expected, the slopes of *ta* and *ws* are positive and the slope of *ha* is negative.

Predictions of NH_3 emissions corresponding to the cross-validation dataset were produced for each of the four variable combinations. The rows labeled "% in PI" and "Width" present diagnostics that describe the quality of uncertainty quantification of the EEM. Ideally, 95 percent of the observed NH_3 emissions in the cross-validation dataset should fall inside the 95 percent prediction intervals. For all of the EEMs, the value exceeds 95 percent which means that the intervals are somewhat conservative. The width of the 95 percent prediction intervals gives some notion of their usefulness. For example, an interval with width 4 kg is more useful than one with width of 7 kg, if both had the same "% in PI."

Table 5-12. Fit Statistics for Subsets of Mean Trend Variables

Fit Statistic	Mean Trend Variables[a]						
	Individual Met Variables			Solar Radiation Variables		All Main Effects	Two-Way Interactions
	ta	*ha*	*ws*	*sh, ch*	*sd, cd*		
Slope	0.53	-0.27	-0.38	NA[b]	NA	NA	NA
% in PI	97	100	96	99	96	99	99
Width (kg)	4.0	7.6	3.9	6.5	3.9	4.3	4.4
R^2_{Test}	0.33	0.06	0.21	0.01	0.27	0.58	0.59
RMSE (kg)	1.2	1.4	1.3	1.4	1.2	0.93	0.91
γ_0 (kg)	-0.32 x	0.10 ✓	0.25 x	0.06 ✓	-0.24 x	-0.11 x	-0.00 ✓
γ_1	1.4 x	0.83 x	0.95 ✓	0.87 ✓	1.4 x	1.2 x	1.0 x[c]

[a] A check mark (✓) in the column for γ_0 indicates that the estimate is not significantly different from zero at the $\alpha = 0.05$ significance level, while an "x" indicates that it is significantly different from zero. Similarly, a check mark or an "x" in the column for γ_1 indicates whether the estimate is significantly different from one.
[b] NA = Not applicable.
[c] This value was significantly different than one when additional decimals places were considered.

The fit statistics shown in the last four rows of Table 5-12 result from a linear regression of measured NH_3 emissions in the cross-validation dataset on the point predictions of NH_3 emissions. The correlation coefficient, R^2, tells what proportion of the variability of emissions in the cross-validation dataset was explained by the EEM. Table 5-12 shows that the EEMs that used *ta*, *ws*, and the *sd*, *cd* combination, respectively, explained the most variability. When all main effects were included, 58 percent of the variability in the cross-validation dataset emissions was explained. When all main effects and two-way interactions were included, 59 percent of the variability was explained. Because the increase in variability explained by progressing from main effects only to including two-way interactions was only 1 percent, the EPA did not consider three-way interactions among meteorological variables.

The root mean squared error (RMSE), in kg NH_3, is defined in Equation 5-7. The RMSE provides a measure of the typical deviation of an NH_3 emissions value in the cross-validation dataset from the point prediction of NH_3 emissions. Thus, smaller values of RMSE indicate a better fit. Notice that of the EEMs including individual predictor variables, those containing *ta* and the *sd*, *cd* combination resulted in the lowest RMSE (1.2 kg). The EEM including all main effects reduced the RMSE to 0.93 kg and including all two-way interactions further reduced the RMSE value to 0.91 kg.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y_{ht} - \hat{Y}_{ht}\right)^{2}} \qquad \text{Equation 5-7}$$

If the EEM fit the cross-validation dataset perfectly, the intercept of this linear regression, Υ_0, would equal zero and the slope, Υ_1, would equal one. If Υ_0 or Υ_1 is significantly different from 0 or 1, respectively, the EEM has systematic bias. A check mark (✓) beside either Υ_0 or Υ_1 indicates that it is not significantly different from zero or one, respectively. Adding the two-way interactions to the main effects resulted in the intercept being closer to 0 and the estimated slope became 1 when rounded to the nearest tenth. Despite the fact that the difference between the intercept and 1 is statistically significant when additional decimal places were considered, the difference does not have practical significance.

5.4 Choosing the Covariance Function

Although the link function is not part of the covariance function, this section begins by showing how the EPA chose the log link using the working mean trend function. After this decision was made, the EPA used an EEM with the log link function to evaluate different covariance function components.

5.4.1 Selecting the link function

As mentioned in Section 5.2, three link functions can be used to relate the mean trend function to the expected value of NH_3 emissions under a given set of conditions: the identity, the natural logarithm, and the reciprocal. If the link function is the identity, then the mean of the distribution is exactly equal to the mean trend function. If the link function is the natural logarithm, then the natural log of the distribution mean is equal to the mean tread function. Finally, if the link function is the reciprocal, then the inverse of the distribution mean is equal to the inverse of the mean trend function.

To compare the three link functions, the EPA fit an EEM with the working mean trend function and each of the link functions, and recorded the fit statistics in Table 5-13. Notice that while the proportions of variability in the cross-validation dataset emissions explained by the identity and log link EEMs were similar, the 95 percent prediction intervals produced when the identity link was used were so narrow as to not contain any of the emissions in the cross-validation dataset. Thus, the quantification of uncertainty resulting from the use of the identity function was inaccurate.

The reciprocal link EEM explains only 48 percent of the variability in emissions of the cross-validation dataset, compared to the 59 percent explained by the log link EEM. The conservative 100 percent coverage probability of the prediction intervals produced by the reciprocal link EEM was achieved by prediction intervals with mean width 5.8 kg. The 99 percent coverage probability of the log link prediction intervals have a narrower, and thus more useful mean width of 4.4 kg.

Table 5-13. Fit Statistics for Identity, Log and Reciprocal Link Functions

Fit Statistic[a]	Link functions		
	Identity	Log	Reciprocal
R^2	0.57	0.59	0.48
% in PI	0	99	100
Mean width PI's (kg)	0.042	4.4	5.8
RMSE (kg)	0.94	0.91	1.0
γ_0 (kg)	0.027 ✓	-0.00 ✓	0.25 x
γ_1	53 x	1.0 ✓	0.61 x

[a] A check mark (✓) in the column for γ_0 indicates that the estimate is not significantly different from zero at the $\alpha = 0.05$ significance level, while an "x" indicates that it is significantly different from zero. Similarly, a check mark or an "x" in the column for γ_1 indicates whether the estimate is significantly different from one.

5.4.2 Serial correlation

A possible source of correlation in the deviations from the mean trend function, e_{st}, is serial correlation, which is usually present in time series data. Failure to account for serial correlation can result in prediction interval widths that are overly optimistic and biased regression coefficient estimates for finite sample sizes. Accounting for serial correlation can be accomplished by: using a function of time in the mean trend function, using a parametric covariance function such as the AR(1) or using a nonparametric covariance structure such as a radial smoother.

Normally, diagnosis of the covariance structure involves plotting lagged residuals as a function of current time residuals, but the large number of large gaps in the hourly NH_3 emissions values made such plots uninformative. The EPA tried a parametric AR(1) structure, but the autoregressive parameter was not statistically significant. The EPA also used a radial smoother as a nonparametric covariance structure, but the numerical algorithm designed to compute basis functions required for such a structure returned an error.

Thus, the EPA used the functions of time already included in the mean trend function to account for relationships between NH_3 at one time and NH_3 at an adjacent time. The diurnal and annual cycles account for two different units of time and they interact with one another and with all other mean trend variables. As the fit statistics in Table 5-13 indicated at this point in the analysis, the prediction interval widths for the EEM with the working mean trend function and the log link account for enough variability for the coverage probability to exceed 95 percent. This result shows that not accounting for serial correlation beyond use of the periodic functions of hour and day in the mean trend did not result in intervals that were too narrow. Therefore, the EPA did not use serial correlation.

5.4.3 Random effect of site

If conditions not captured by the mean trend function cause different sites to have different NH_3 emissions, a random effect of site might account for the additional prediction uncertainty caused by this site-specific variability. A fixed effect of site, as opposed to a random effect, is not appropriate because the EEM will be used to estimate emissions from sites that did not participate in the NAEMS.

The difficulty of including a random effect is that while it is not part of the mean trend function, it would have to replace the use of one of the static farm-based predictor variables. Site is a static farm-based variable also, and its use as a random effect would add another parameter to the EEM, using up one of the degrees of freedom available to learn about the static farm-based variables.

Despite this difficulty, the EPA examined fit statistics resulting from the inclusion of a random effect in the EEM that used the log link and the working mean trend function. The EPA fit an EEM with a random effect of site to the base dataset, and then tested the null hypothesis that the variance component associated with a random effect of site, σ_S^2, was equal to zero using a Wald test and a chi-squared test. For the Wald test, the p-value was an unconvincing 0.06. The chi-squared test produced a p-value of 0.0001, which seemed to suggest that a random effect of site allowed the EEM to better fit the base dataset. This EEM, however, does not contain another farm-based variable to describe site-to-site variability. Consequently, at this phase of the analysis, the random effect of site was accounting for any site-to-site differences, regardless of their source.

Table 5-14 gives the fit statistics that describe how well the EEM with a random effect of site fit the cross-validation dataset. Including the random effect in the EEM resulted in a coverage probability that was further away from 95 percent than that achieved by the EEM that did not include a random effect. Also, the mean width of the prediction intervals increased by 3.1 kg. The percent of variability in emissions in the cross-validation dataset that was explained increased from 59 percent to 60 percent upon inclusion of the random effect and the RMSE value decreased. The intercept of the regression of cross-validation emissions on the point predictions moved further from 0 to become significantly different from 0. Additionally, the slope moved further from 1. Because most of the fit statistics for the cross-validation dataset did not improve when the random effect of site was added to the EEM, the EPA concluded that this version of the EEM over-fit the base dataset.

As a final test of the usefulness of including a random effect of site in the EEM, the EPA progressively added each of the three chosen farm-based predictor variables to the working mean trend function of the random effect-EEM. Even when only one farm-based predictor variable was added at a time to the EEM, the fitting algorithm did not converge. This result suggested that when a static, farm-based predictor variable that was expected to inform variability in NH_3 emissions was included in the EEM, the random effect of site provided redundant information. Consequently, the fitting algorithm oscillated between different optimal values of the parameters.

Based on the poor fit to the cross-validation dataset and the lack of convergence of the fitting algorithm, the EPA did not consider a random effect of site in development of the EEM.

Table 5-14. Fit Statistics With and Without a Random Effect of Site

Fit Statistic[a]	Random Effect of Site Included in EEM?	
	No	Yes
% in PI	99	100
Mean width PI's (kg)	4.4	7.5
R^2	0.59	0.60
RMSE (kg)	0.91	0.90
γ_0 (kg)	-0.00 ✓	-0.29 x
γ_1	1.0 ✓	1.5 x

[a] A check mark (✓) in the column for γ_0 indicates that the estimate is not significantly different from zero at the $\alpha = 0.05$ significance level, while an "x" indicates that it is significantly different from zero. Similarly, a check mark or an "x" in the column for γ_1 indicates whether the estimate is significantly different from one

5.5 Comparison of EEMs With Different Mean Trend Variables

To choose the final EEM, the EPA evaluated the performance of three EEMs. Each of the EEMs contained two of the three candidate farm-based predictor variable (i.e., *animal*, *sa* and *size*), and the interactions between those variable pairs and the main effects and 2-way interactions of the meteorological and time-based variables. Table 5-15 shows the fit statistics calculated for each EEM on the cross-validation dataset. All three versions of the EEM had a coverage probability of 99 percent, and while the prediction interval widths differ by at most 0.1 kg, the optimal width of 4.5 kg is that of the *animal/size* EEM. The *animal/sa* EEM explained the most variability in cross-validation emissions, at 74 percent, with a large gap separating it from the other two EEMs. The *sa/size* EEM is the second best at 68 percent, and the *animal/size* EEM explained 66 percent. The optimal value of RMSE, 0.73 kg, was again achieved by the *animal/sa* EEM, and again a large gap separates it from the others. The *sa/size* EEM again has the second best value, at 0.80 kg, followed by the *animal/size* EEM at 0.83 kg.

For the *animal/sa* EEM, neither the slope nor the intercept of the regression of cross-validation NH_3 emissions on predictions of them was significantly different from the optimal values of 0 and 1, respectively. For the other two EEMs, both the slope and intercept are significantly different from the optimal values. Based on these fit statistics, it seems clear that of these three EEMs, the one that best fits the data is the *animal/sa* EEM.

Table 5-15. Fit Statistics for Three Combinations of Two Farm-Based Variables

Fit statistic	Farm-Based Predictor Variable Pair		
	animal/sa	***animal/size***	***sa/size***
% in PI	99	99	99
Width (kg)	4.6	4.5	4.6
R^2	0.74	0.66	0.68
RMSE (kg)	0.73	0.83	0.80
γ_0 (kg)	-0.00 ✓	-0.11 x	0.11 x
γ_1	1.0 ✓	1.1 x	0.92 x

[a] A check mark (✓) in the column for γ_0 indicates that the estimate is not significantly different from zero at the α = 0.05 significance level, while an “x” indicates that it is significantly different from zero. Similarly, a check mark or an “x” in the column for γ_1 indicates whether the estimate is significantly different from one

5.6 NH_3 EEM Summary, Use and Implications

This section presents the calculations using the three candidate EEMs presented in Section 5.5 for several example swine and dairy farms. For this analysis, the parameter estimates for each EEM were obtained from fitting each to the full dataset (i.e., combined base and cross-validation datasets).

Table 5-16 shows the values of the example predictor variables and the corresponding mean trend variables. The values of the mean trend variables were obtained from the values of the predictor variables as follows.

- *sd* and *cd*: These values are the sinusoidal periodic functions of Julian day, as defined in Equation 5-2, for this example, *jday* is equal to 243, co *sd* = -0.86169 and *cd* = -0.50743.
- *sh* and *ch*: These values are the sinusoidal periodic functions of hour of the day, as defined in Equation 5-3. For this example, *hour* is equal to for 6:00 p.m., which corresponds to *sh* = -1 and *ch* = -1.84E-16 .
- *ta, ha, ws, sa* and *size*: For these variables, the value of the predictor variable was centered and scaled by subtracting from it the “centering value” and dividing by the “scaling value” (Table 5-9 presents the centering and scaling values).

Table 5-16. Variables Values Used In Example Calculations

Predictor Variable			Mean Trend Variable	
Description	**Value**	**Unit**	**Description**	**Value**
Julian day (*jday*)	243	day	*sd*	-0.86
			cd	-0.51
Hour of the day (*hour*)	18	hour	*sh*	-1
			ch	0
Temperature (*ta**)	29	°C	*ta*	1.3
Humidity (*ha**)	40	%	*ha*	-1.0
Wind speed (*ws**)	4.1	meters/second	*ws*	0.41
Animal type (*animal*)	1	Dairy	*animal*	1
	-1	Swine		-1
Surface area (*sa**)	121	1,000 ft^2	*sa*	-0.44
	200			0.45
Animal capacity * adult weigh (*size**)	1.9	million lb	*size*	-0.042
	0.68			-0.703

[a] An asterisk (*) is used to note that these predictor variables are the original values reported in the data submitted to the EPA before centering and scaling.

To obtain the point estimate (i.e., mean) for each of the three EEMs, the values of the mean trend variables and the estimated regression coefficients from Table 5-18, Table 5-20 or Table 5-22, respectively, were inserted into Equation 5-8.

$$\hat{Y} = e^{(\widehat{\beta_0}+\widehat{\beta_1}x_1+\cdots+\widehat{\beta_{80}}x_{80})} \qquad \text{Equation 5-8}$$

Equation 5-8 estimates the kilograms of NH_3 emitted from the lagoon during a half-hour period. The lower and upper bounds of the 95-percent prediction intervals were calculated by subtracting and adding the margin of error to the point estimate. Because the gamma probability distribution (see Figure 5-5) is not symmetric, the error margins that are added and subtracted from the mean are not equal.

5.6.1 Animal/surface area EEM

For the *animal/sa* EEM examples, the EPA chose lagoon surface areas of 121,000 $ft^2$and 200,000 ft^2. The smaller value of surface area is similar to swine sites OK3A and IN4A and the observed lagoons at IN5A, WA5A and WI5A. The larger surface area is representative of larger swine facilities similar to sites NC3A and NC4A and all the lagoons in a multi-stage system prevalent at dairy facilities.

Table 5-17 shows the point prediction estimates and the lower and upper bounds of the 95-percent prediction interval for each lagoon size at a swine and dairy farm. Table 5-18 presents the mean trend variables (x_{pst}), the values of the regression parameter estimates ($\widehat{\beta_p}$) and the values of each mean trend variable (x_p) for the four example farms.

Table 5-17. Results of the *animal/sa* EEM Examples

Example Number	Farm-Based Predictor Variables		Point Prediction, $\hat{\mu}$ (kg/30 min)	95% Prediction Interval Bounds (kg/30 min)	
	animal[a]	***sa**** (1,000 ft^2) [b]	$\hat{Y}$	Lower	Upper
1	-1	121	1.8	0.05	6.5
2	1	121	1.2	0.04	4.3
3	-1	200	2.2	0.07	7.8
4	1	200	1.4	0.04	5.1

[a] For the variable *animal*, 1 is dairy and -1 is swine.
[b] An asterisk (*) is used to note that these predictor variables are the original values reported in the data submitted to the EPA before centering and scaling.

While the fact that this prediction interval has width 6.5 – 0.55 = 5.95 kg may seem to render it not very useful, the quantities most of interest to the EPA are sums of NH_3 emissions. If individual half-hour emissions follow a gamma distribution, then according to the Central Limit Theorem, the sum of many half-hour emissions will follow a normal distribution. The mean of the normal distribution will be the sum of the means of the half-hour emissions. The variance of the normal distribution will be the sums of the variances of the half-hour emissions. The variance for the gamma distribution is $\mu^2\phi$. For the example above, the point prediction is the estimate of μ, and the estimate of ϕ is given in the last row of Table 5-18. The variance is thus 3.5 kg^2. As a simple example to understand how useful the prediction intervals for a sum would be, imagine 100 half-hour periods just like the one in Example 1. The sum of emissions for all 100 half-hours would have a normal distribution with approximate mean and variance of 180 kg and 350 kg^2, respectively. The point estimate for the total emissions for the 100 half-hour periods would be the mean, 180 kg. The 95 percent prediction interval would be obtained by adding and subtracting 1.96 multiplied by the standard deviation, which is the square root of the variance: $180 \pm 1.96\sqrt{350}$. Thus, the lower bound would be 143 kg and the upper bound would be 217 kg, and this prediction interval would have the same interpretation given above.

For the *animal/sa* EEM, the point predictions for swine are higher than for dairy at both example lagoon sizes. For Example 1 (swine lagoon), the mean emissions rate was 1.8 kg/30 minutes while the mean emissions rate for Example 2 (dairy lagoon) was 1.2 kg/30 minutes. Thus when meteorology and surface area are held constant, this EEM predicts that swine lagoons emit more NH_3. Examples 1 and 3, both represent swine facilities. The point prediction for a surface area of 121,000 ft^2 is 1.8 kg, and for 200,000 ft^2 is 2.2 kg. These predictions of NH_3 emissions are consistent with the expectation that greater surface area presents more opportunities for desorption of NH_3 from the lagoon liquid.

Table 5-18. Values of Mean Trend Variables for the *animal/sa* EEM Examples

p	Name of x_{pst}	$\hat{\beta}_p$	Value of x_p			
			Ex. 1	Ex. 2	Ex. 3	Ex. 4
0	Intercept	-0.1704	NA [a]	NA	NA	NA
1	*animal*	-0.1264	-1	1	-1	1
2	*sa*	0.166	-0.4382	-0.4382	0.4494	0.4494
3	*ta*	0.2238	1.30	1.30	1.30	1.30
4	*ha*	0.01534	-1.00	-1.00	-1.00	-1.00
5	*ws*	0.3378	0.41	0.41	0.41	0.41
6	*sh*	0.06782	-1	-1	-1	-1
7	*ch*	0.01614	-1.84E-16	-1.84E-16	-1.84E-16	-1.84E-16
8	*sd*	-0.1347	-0.86169	-0.86169	-0.86169	-0.86169
9	*cd*	-0.3046	-0.50743	-0.50743	-0.50743	-0.50743
10	*animal*ta*	0.0441	-1.3000	1.3000	-1.3000	1.3000
11	*animal*ha*	0.0528	1.0000	-1.0000	1.0000	-1.0000
12	*animal*ws*	0.03723	-0.4100	0.4100	-0.4100	0.4100
13	*animal*sh*	-0.08584	1.0000	-1.0000	1.0000	-1.0000
14	*animal*ch*	-0.0158	0.0000	0.0000	0.0000	0.0000
15	*animal*sd*	0.09038	0.8617	-0.8617	0.8617	-0.8617
16	*animal*cd*	0.2183	0.5074	-0.5074	0.5074	-0.5074
17	*sa*ta*	0.03896	-0.5697	-0.5697	0.5843	0.5843
18	*sa*ha*	-0.00087	0.4382	0.4382	-0.4494	-0.4494
19	*sa*ws*	0.1086	-0.1797	-0.1797	0.1843	0.1843
20	*sa*sh*	0.04138	0.4382	0.4382	-0.4494	-0.4494
21	*sa*ch*	-0.01069	0.0000	0.0000	0.0000	0.0000
22	*sa*sd*	0.03115	0.3776	0.3776	-0.3873	-0.3873
23	*sa*cd*	-0.04754	0.2224	0.2224	-0.2281	-0.2281
24	*ta*ha*	0.05177	-1.3000	-1.3000	-1.3000	-1.3000
25	*ta*ws*	0.04942	0.5330	0.5330	0.5330	0.5330
26	*ta*sh*	-0.05717	-1.3000	-1.3000	-1.3000	-1.3000
27	*ta*ch*	0.08774	0.0000	0.0000	0.0000	0.0000
28	*ta*sd*	-0.0417	-1.1202	-1.1202	-1.1202	-1.1202
29	*ta*cd*	-0.05216	-0.6597	-0.6597	-0.6597	-0.6597
30	*ha*ws*	0.07141	-0.4100	-0.4100	-0.4100	-0.4100
31	*ha*sh*	-0.02224	1.0000	1.0000	1.0000	1.0000
32	*ha*ch*	-0.00842	0.0000	0.0000	0.0000	0.0000
33	*ha*sd*	-0.04523	0.8617	0.8617	0.8617	0.8617
34	*ha*cd*	-0.06174	0.5074	0.5074	0.5074	0.5074
35	*ws*sh*	-0.04305	-0.4100	-0.4100	-0.4100	-0.4100

Table 5-18. Values of Mean Trend Variables for the *animal/sa* EEM Examples

p	Name of x_{pst}	$\hat{\beta}_p$	Value of x_p			
			Ex. 1	Ex. 2	Ex. 3	Ex. 4
36	*ws*ch*	-0.00906	0.0000	0.0000	0.0000	0.0000
37	*ws*sd*	-0.02067	-0.3533	-0.3533	-0.3533	-0.3533
38	*ws*cd*	0.01939	-0.2080	-0.2080	-0.2080	-0.2080
39	*sh*sd*	0.003985	0.8617	0.8617	0.8617	0.8617
40	*sh*cd*	0.01495	0.5074	0.5074	0.5074	0.5074
41	*ch*sd*	0.03812	1.58E-16	1.58E-16	1.58E-16	1.58E-16
42	*ch*cd*	0.07877	9.32E-17	9.32E-17	9.32E-17	9.32E-17
43	*animal*ta*ha*	-0.0058	1.30	-1.30	1.30	-1.30
44	*animal*ta*ws*	0.009751	-0.53	0.53	-0.53	0.53
45	*animal*ta*sh*	0.006322	1.30	-1.30	1.30	-1.30
46	*animal*ta*ch*	-0.03063	0.00	0.00	0.00	0.00
47	*animal*ta*sd*	0.07724	1.12	-1.12	1.12	-1.12
48	*animal*ta*cd*	-0.01519	0.66	-0.66	0.66	-0.66
49	*animal*ha*ws*	0.01062	0.41	-0.41	0.41	-0.41
50	*animal*ha*sh*	0.06862	-1.00	1.00	-1.00	1.00
51	*animal*ha*ch*	-0.01994	0.00	0.00	0.00	0.00
52	*animal*ha*sd*	-0.01027	-0.86	0.86	-0.86	0.86
53	*animal*ha*cd*	-0.1979	-0.51	0.51	-0.51	0.51
54	*animal*ws*sh*	0.01821	0.41	-0.41	0.41	-0.41
55	*animal*ws*ch*	0.01005	0.00	0.00	0.00	0.00
56	*animal*ws*sd*	-0.004	0.35	-0.35	0.35	-0.35
57	*animal*ws*cd*	0.022	0.21	-0.21	0.21	-0.21
58	*animal*sh*sd*	0.07842	-0.86	0.86	-0.86	0.86
59	*animal*sh*cd*	0.1134	-0.51	0.51	-0.51	0.51
60	*animal*ch*sd*	-0.05731	0.00	0.00	0.00	0.00
61	*animal*ch*cd*	-0.07813	0.00	0.00	0.00	0.00
62	*sa*ta*ha*	-0.01088	0.57	0.57	-0.58	-0.58
63	*sa*ta*ws*	0.005086	-0.23	-0.23	0.24	0.24
64	*sa*ta*sh*	-0.03341	0.57	0.57	-0.58	-0.58
65	*sa*ta*ch*	0.02372	0.00	0.00	0.00	0.00
66	*sa*ta*sd*	0.05193	0.49	0.49	-0.50	-0.50
67	*sa*ta*cd*	-0.04981	0.29	0.29	-0.30	-0.30
68	*sa*ha*ws*	0.03486	0.18	0.18	-0.18	-0.18
69	*sa*ha*sh*	0.0476	-0.44	-0.44	0.45	0.45
70	*sa*ha*ch*	-0.01861	0.00	0.00	0.00	0.00
71	*sa*ha*sd*	0.06239	-0.38	-0.38	0.39	0.39
72	*sa*ha*cd*	-0.08068	-0.22	-0.22	0.23	0.23
73	*sa*ws*sh*	0.01293	0.18	0.18	-0.18	-0.18
74	*sa*ws*ch*	0.01629	0.00	0.00	0.00	0.00
75	*sa*ws*sd*	0.02821	0.15	0.15	-0.16	-0.16
76	*sa*ws*cd*	-0.02738	0.09	0.09	-0.09	-0.09
77	*sa*sh*sd*	-0.03134	-0.38	-0.38	0.39	0.39
78	*sa*sh*cd*	-0.1219	-0.22	-0.22	0.23	0.23
79	*sa*ch*sd*	0.008813	-6.94E-17	-6.94E-17	7.11E-17	7.11E-17

Table 5-18. Values of Mean Trend Variables for the *animal/sa* EEM Examples

p	Name of x_{pst}	$\hat{\beta}_p$	Value of x_p			
			Ex. 1	Ex. 2	Ex. 3	Ex. 4
80	*sa*ch*cd*	-0.00419	-4.08E-17	-4.08E-17	4.19E-17	4.19E-17
	Scale parameter	1.061	NA	NA	NA	NA

[a] NA = Not applicable.

5.6.2 Animal/size EEM

For the *animal/size* EEM examples, the EPA chose values of farm size that represented plausible combinations of *size** for both swine and dairy facilities, as supported by the USDA Agricultural Census. The first value of *size** (i.e., *capacity** multiplied by *adultwt**) chosen by the EPA was 1.9 million pounds. Such a value would occur at a dairy facility slightly smaller than that at site WI5A, with total capacity equal to 1,357 head of milk cows at 1,400 pounds each. This value of *size** would also occur at a swine breeding-gestation farm slightly larger than site OK4A, with a *capacity** of 3,878 sows at 490 pounds each for the same value of *adultwt**. The second value of *size** selected was 0.68 million pounds. This value would occur at a dairy site with a total capacity of 486 head (similar to site WI5A); or at a swine site with 1,388 pigs (similar to sites IA3A and IN4A) for the same value of *adultwt**. This value of *size** is also representative of swine operations similar to those at sites IA3A and IN4A.

The point prediction estimates and the lower and upper bounds of the 95-percent prediction interval for each farm size example at both a swine and a dairy operation are shown in Table 5-19. Table 5-20 presents the mean trend variables (x_{pst}), the values of the regression parameter estimates ($\widehat{\beta_p}$) and the values of each mean trend variable for the four examples

Table 5-19. Results of the *animal/size* EEM Examples

Example Number	Farm-Based Predictor Variables		Point Prediction (kg/30 min)	95% Prediction Interval Bounds (kg/30 min)	
	animal[a]	*size** (lbs) [b]	$\hat{Y}$	Lower	Upper
1	-1	1,900,000	2.4	0.07	8.6
2	1	1,900,000	0.95	0.03	3.4
3	-1	680,000	1.9	0.06	6.9
4	1	680,000	0.8	0.02	2.8

[a] For the variable *animal*, 1 is dairy and -1 is swine.
[b] An asterisk (*) is used to note that these predictor variables are the original values reported in the data submitted to the EPA before centering and scaling.

The point prediction for the swine site is 2.4 kg/30 minutes, while that for the dairy site is 0.95 kg/30 minutes. These predictions of NH_3 emissions are comparable to the daily emissions values contained in the final reports (see Section 3).

For the dairy Examples 1 and 3, the point prediction when *size** = 1.9 million pounds is 2.4 kg/30 minutes, and the point prediction when *size** = 0.68 million pounds is 1.9 kg/30 minutes. The swine emission pattern is similar. These predictions of NH_3 emissions are consistent with the expectation that greater live animal mass should result in more precursors supplied to the lagoon, which should result in more NH_3 emissions.

Table 5-20. Values of Mean Trend Variables for *animal/size* EEM Examples

p	Name of x_{pst}	$\hat{\beta}_p$	Value of x_p			
			Ex. 1	Ex. 2	Ex. 3	Ex. 4
0	Intercept	-0.3005	NA [a]	NA	NA	NA
1	*animal*	-0.5471	-1	1	-1	1
2	*size*	0.4147	-0.0422	-0.0422	-0.7033	-0.7033
3	*ta*	0.1932	1.30	1.30	1.30	1.30
4	*ha*	-0.00455	-1.00	-1.00	-1.00	-1.00
5	*ws*	0.2161	0.41	0.41	0.41	0.41
6	*sh*	0.05094	-1	-1	-1	-1
7	*ch*	0.04161	-1.84E-16	-1.84E-16	-1.84E-16	-1.84E-16
8	*sd*	-0.2406	-0.86169	-0.86169	-0.86169	-0.86169
9	*cd*	-0.3591	-0.50743	-0.50743	-0.50743	-0.50743
10	*animal*ta*	-0.2149	-1.3000	1.3000	-1.3000	1.3000
11	*animal*ha*	0.000272	1.0000	-1.0000	1.0000	-1.0000
12	*animal*ws*	-0.3289	-0.4100	0.4100	-0.4100	0.4100
13	*animal*sh*	-0.2317	1.0000	-1.0000	1.0000	-1.0000
14	*animal*ch*	0.08868	0.0000	0.0000	0.0000	0.0000
15	*animal*sd*	-0.287	0.8617	-0.8617	0.8617	-0.8617
16	*animal*cd*	-0.1205	0.5074	-0.5074	0.5074	-0.5074
17	*size*ta*	0.2064	-0.0548	-0.0548	-0.9143	-0.9143
18	*size*ha*	0.02935	0.0422	0.0422	0.7033	0.7033
19	*size*ws*	0.4158	-0.0173	-0.0173	-0.2884	-0.2884
20	*size*sh*	0.1528	0.0422	0.0422	0.7033	0.7033
21	*size*ch*	-0.1077	0.0000	0.0000	0.0000	0.0000
22	*size*sd*	0.3486	0.0364	0.0364	0.6060	0.6060
23	*size*cd*	0.2875	0.0214	0.0214	0.3569	0.3569
24	*ta*ha*	0.04663	-1.3000	-1.3000	-1.3000	-1.3000
25	*ta*ws*	0.007551	0.5330	0.5330	0.5330	0.5330
26	*ta*sh*	-0.01768	-1.3000	-1.3000	-1.3000	-1.3000
27	*ta*ch*	0.1057	0.0000	0.0000	0.0000	0.0000
28	*ta*sd*	-0.07326	-1.1202	-1.1202	-1.1202	-1.1202
29	*ta*cd*	0.01755	-0.6597	-0.6597	-0.6597	-0.6597
30	*ha*ws*	0.005272	-0.4100	-0.4100	-0.4100	-0.4100
31	*ha*sh*	-0.06899	1.0000	1.0000	1.0000	1.0000

Table 5-20. Values of Mean Trend Variables for *animal/size* EEM Examples

p	Name of x_{pst}	$\hat{\beta}_p$	Value of x_p			
			Ex. 1	Ex. 2	Ex. 3	Ex. 4
32	*ha*ch*	-0.00278	0.0000	0.0000	0.0000	0.0000
33	*ha*sd*	-0.04194	0.8617	0.8617	0.8617	0.8617
34	*ha*cd*	0.0573	0.5074	0.5074	0.5074	0.5074
35	*ws*sh*	-0.0395	-0.4100	-0.4100	-0.4100	-0.4100
36	*ws*ch*	-0.02627	0.0000	0.0000	0.0000	0.0000
37	*ws*sd*	-0.08146	-0.3533	-0.3533	-0.3533	-0.3533
38	*ws*cd*	-0.01699	-0.2080	-0.2080	-0.2080	-0.2080
39	*sh*sd*	0.003667	0.8617	0.8617	0.8617	0.8617
40	*sh*cd*	0.06681	0.5074	0.5074	0.5074	0.5074
41	*ch*sd*	0.08076	1.58E-16	1.58E-16	1.58E-16	1.58E-16
42	*ch*cd*	0.1565	9.32E-17	9.32E-17	9.32E-17	9.32E-17
43	*animal*ta*ha*	-0.08574	1.3000	-1.3000	1.3000	-1.3000
44	*animal*ta*ws*	-0.1486	-0.5330	0.5330	-0.5330	0.5330
45	*animal*ta*sh*	0.05419	1.3000	-1.3000	1.3000	-1.3000
46	*animal*ta*ch*	0.02732	0.0000	0.0000	0.0000	0.0000
47	*animal*ta*sd*	-0.03565	1.1202	-1.1202	1.1202	-1.1202
48	*animal*ta*cd*	0.3233	0.6597	-0.6597	0.6597	-0.6597
49	*animal*ha*ws*	-0.146	0.4100	-0.4100	0.4100	-0.4100
50	*animal*ha*sh*	-0.04517	-1.0000	1.0000	-1.0000	1.0000
51	*animal*ha*ch*	-0.00081	-1.84E-16	1.84E-16	-1.84E-16	1.84E-16
52	*animal*ha*sd*	-0.1607	-0.8617	0.8617	-0.8617	0.8617
53	*animal*ha*cd*	0.04089	-0.5074	0.5074	-0.5074	0.5074
54	*animal*ws*sh*	0.009941	0.4100	-0.4100	0.4100	-0.4100
55	*animal*ws*ch*	-0.00194	7.53E-17	-7.53E-17	7.53E-17	-7.53E-17
56	*animal*ws*sd*	-0.2062	0.3533	-0.3533	0.3533	-0.3533
57	*animal*ws*cd*	-0.1885	0.2080	-0.2080	0.2080	-0.2080
58	*animal*sh*sd*	0.06039	-0.8617	0.8617	-0.8617	0.8617
59	*animal*sh*cd*	0.2116	-0.5074	0.5074	-0.5074	0.5074
60	*animal*ch*sd*	0.005975	-1.58E-16	1.58E-16	-1.58E-16	1.58E-16
61	*animal*ch*cd*	0.04529	-9.32E-17	9.32E-17	-9.32E-17	9.32E-17
62	*size*ta*ha*	0.09349	0.0548	0.0548	0.9143	0.9143
63	*size*ta*ws*	0.2109	-0.0225	-0.0225	-0.3749	-0.3749
64	*size*ta*sh*	-0.04622	0.0548	0.0548	0.9143	0.9143
65	*size*ta*ch*	-0.06496	1.01E-17	1.01E-17	1.68E-16	1.68E-16
66	*size*ta*sd*	0.06439	0.0473	0.0473	0.7878	0.7878
67	*size*ta*cd*	-0.3927	0.0278	0.0278	0.4639	0.4639
68	*size*ha*ws*	0.1765	0.0173	0.0173	0.2884	0.2884
69	*size*ha*sh*	0.1073	-0.0422	-0.0422	-0.7033	-0.7033
70	*size*ha*ch*	-0.02489	-7.75E-18	-7.75E-18	-1.29E-16	-1.29E-16
71	*size*ha*sd*	0.1429	-0.0364	-0.0364	-0.6060	-0.6060
72	*size*ha*cd*	-0.2447	-0.0214	-0.0214	-0.3569	-0.3569
73	*size*ws*sh*	0.01065	0.0173	0.0173	0.2884	0.2884
74	*size*ws*ch*	0.008559	3.18E-18	3.18E-18	5.30E-17	5.30E-17
75	*size*ws*sd*	0.2204	0.0149	0.0149	0.2485	0.2485

Table 5-20. Values of Mean Trend Variables for *animal/size* EEM Examples

p	Name of x_{pst}	$\hat{\beta}_p$	Value of x_p			
			Ex. 1	Ex. 2	Ex. 3	Ex. 4
76	*size*ws*cd*	0.3026	0.0088	0.0088	0.1463	0.1463
77	*size*sh*sd*	0.009514	-0.0364	-0.0364	-0.6060	-0.6060
78	*size*sh*cd*	-0.0549	-0.0214	-0.0214	-0.3569	-0.3569
79	*size*ch*sd*	-0.071	-6.68E-18	-6.68E-18	-1.11E-16	-1.11E-16
80	*size*ch*cd*	-0.1125	-3.93E-18	-3.93E-18	-6.56E-17	-6.56E-17
	Scale parameter	1.061	NA	NA	NA	NA

[a] NA = Not applicable.

5.6.3 Surface area/size EEM

The *sa/size* EEM estimates emissions based on farm size and surface area irrespective of animal type. For the *sa/size* EEM, the four examples use the same parameter values as the previous two examples: surface areas of 121,000 ft^2 and 680,000 ft^2 with a farm size of 1,900,000 pounds and 680,000 pounds for each surface area. The point prediction estimates and the lower and upper bounds of the 95-percent prediction interval for each example are shown in Table 5-21. Table 5-22 presents the mean trend variables (x_{pst}), the values of the regression parameter estimates ($\widehat{\beta_p}$) and the values of each mean trend variable for the four examples.

Table 5-21. Results of the *sa/size* EEM Examples

Example Number	Farm-Based Predictor Variables [a]		Point Prediction (kg/30 min)	95% Prediction Interval Bounds (kg/30 min)	
	*sa** (1,000 ft^2)	*size** (lbs)	$\hat{Y}$	Lower	Upper
1	121	1,900,000	1.6	0.05	5.8
2	121	680,000	1.8	0.05	6.5
3	200	1,900,000	2.0	0.06	7.2
4	200	680,000	2.2	0.07	8.0

[a] An asterisk (*) is used to note that these predictor variables are the original values reported in the data submitted to the EPA before centering and scaling.

For the same farm size, the point prediction for surface area of 121,000 ft^2 is 1.6 kg/30 minutes, and the prediction for surface area of 200,000 ft^2 is 2.0 kg/30 minutes. These predictions of NH_3 emissions are consistent with the expectation that greater surface area allows for greater desorption of NH_3 from the lagoon liquid.

For the same surface area, the point prediction for farm size of 1.9 million pounds is 1.6 kg/30 minutes, and the point prediction farm size is 680,000 pounds is 1.8 kg/30 minutes.

These predictions of NH_3 emissions are not consistent with the expectation that greater live animal mass should result in more precursors supplied to the lagoon, which, in turn should result in more NH_3 emissions.

Table 5-22. Values of Mean Trend Variables for the *sa/size* EEM Examples

p	x_{pst}	$\hat{\beta}_p$	**Value of x_p**			
			Ex. 1	**Ex. 2**	**Ex. 3**	**Ex. 4**
0	Intercept	-0.1122	NA[a]	NA	NA	NA
1	*size*	-0.09511	-0.0422	-0.7033	-0.0422	-0.7033
2	*sa*	0.2358	-0.4382	-0.4382	0.4494	0.4494
3	*ta*	0.2059	1.30	1.30	1.30	1.30
4	*ha*	-0.02163	-1.00	-1.00	-1.00	-1.00
5	*ws*	0.3208	0.41	0.41	0.41	0.41
6	*sh*	0.1078	-1	-1	-1	-1
7	*ch*	0.01653	-1.84E-16	-1.84E-16	-1.84E-16	-1.84E-16
8	*sd*	-0.1791	-0.8617	-0.8617	-0.8617	-0.8617
9	*cd*	-0.4074	-0.5074	-0.5074	-0.5074	-0.5074
10	*size*ta*	0.05834	-0.0548	-0.9143	-0.0548	-0.9143
11	*size*ha*	0.01868	0.0422	0.7033	0.0422	0.7033
12	*size*ws*	0.03671	-0.0173	-0.2884	-0.0173	-0.2884
13	*size*sh*	-0.05563	0.0422	0.7033	0.0422	0.7033
14	*size*ch*	-0.02925	7.75E-18	1.29E-16	7.75E-18	1.29E-16
15	*size*sd*	0.0874	0.0364	0.6060	0.0364	0.6060
16	*size*cd*	0.2198	0.0214	0.3569	0.0214	0.3569
17	*sa*ta*	0.01243	-0.5697	-0.5697	0.5843	0.5843
18	*sa*ha*	-0.04419	0.4382	0.4382	-0.4494	-0.4494
19	*sa*ws*	0.09624	-0.1797	-0.1797	0.1843	0.1843
20	*sa*sh*	0.08891	0.4382	0.4382	-0.4494	-0.4494
21	*sa*ch*	-0.01693	8.05E-17	8.05E-17	-8.26E-17	-8.26E-17
22	*sa*sd*	-0.00556	0.3776	0.3776	-0.3873	-0.3873
23	*sa*cd*	-0.1107	0.2224	0.2224	-0.2281	-0.2281
24	*ta*ha*	0.06729	-1.3000	-1.3000	-1.3000	-1.3000
25	*ta*ws*	0.04473	0.5330	0.5330	0.5330	0.5330
26	*ta*sh*	-0.071	-1.3000	-1.3000	-1.3000	-1.3000
27	*ta*ch*	0.1068	-2.39E-16	-2.39E-16	-2.39E-16	-2.39E-16
28	*ta*sd*	-0.06093	-1.1202	-1.1202	-1.1202	-1.1202
29	*ta*cd*	-0.0533	-0.6597	-0.6597	-0.6597	-0.6597
30	*ha*ws*	0.06418	-0.4100	-0.4100	-0.4100	-0.4100
31	*ha*sh*	-0.05328	1.0000	1.0000	1.0000	1.0000
32	*ha*ch*	-0.00199	1.84E-16	1.84E-16	1.84E-16	1.84E-16
33	*ha*sd*	-0.01954	0.8617	0.8617	0.8617	0.8617
34	*ha*cd*	0.04567	0.5074	0.5074	0.5074	0.5074
35	*ws*sh*	-0.05145	-0.4100	-0.4100	-0.4100	-0.4100
36	*ws*ch*	-0.01092	-7.53E-17	-7.53E-17	-7.53E-17	-7.53E-17
37	*ws*sd*	-0.01849	-0.3533	-0.3533	-0.3533	-0.3533
38	*ws*cd*	0.02467	-0.2080	-0.2080	-0.2080	-0.2080
39	*sh*sd*	-0.03892	0.8617	0.8617	0.8617	0.8617
40	*sh*cd*	-0.05384	0.5074	0.5074	0.5074	0.5074
41	*ch*sd*	0.0603	1.58E-16	1.58E-16	1.58E-16	1.58E-16
42	*ch*cd*	0.1102	9.32E-17	9.32E-17	9.32E-17	9.32E-17
43	*size*ta*ha*	0.004222	0.0548	0.9143	0.0548	0.9143

Table 5-22. Values of Mean Trend Variables for the *sa/size* EEM Examples

p	x_{pst}	$\hat{\beta}_p$	Value of x_p			
			Ex. 1	**Ex. 2**	**Ex. 3**	**Ex. 4**
44	*size*ta*ws*	0.02358	-0.0225	-0.3749	-0.0225	-0.3749
45	*size*ta*sh*	0.008542	0.0548	0.9143	0.0548	0.9143
46	*size*ta*ch*	-0.0444	1.01E-17	1.68E-16	1.01E-17	1.68E-16
47	*size*ta*sd*	0.07121	0.0473	0.7878	0.0473	0.7878
48	*size*ta*cd*	-0.04238	0.0278	0.4639	0.0278	0.4639
49	*size*ha*ws*	0.01138	0.0173	0.2884	0.0173	0.2884
50	*size*ha*sh*	0.04908	-0.0422	-0.7033	-0.0422	-0.7033
51	*size*ha*ch*	-0.01311	-7.75E-18	-1.29E-16	-7.75E-18	-1.29E-16
52	*size*ha*sd*	0.006304	-0.0364	-0.6060	-0.0364	-0.6060
53	*size*ha*cd*	-0.1493	-0.0214	-0.3569	-0.0214	-0.3569
54	*size*ws*sh*	0.01724	0.0173	0.2884	0.0173	0.2884
55	*size*ws*ch*	0.009766	3.18E-18	5.30E-17	3.18E-18	5.30E-17
56	*size*ws*sd*	0.005532	0.0149	0.2485	0.0149	0.2485
57	*size*ws*cd*	0.07122	0.0088	0.1463	0.0088	0.1463
58	*size*sh*sd*	0.06257	-0.0364	-0.6060	-0.0364	-0.6060
59	*size*sh*cd*	0.09059	-0.0214	-0.3569	-0.0214	-0.3569
60	*size*ch*sd*	-0.0596	-6.68E-18	-1.11E-16	-6.68E-18	-1.11E-16
61	*size*ch*cd*	-0.08964	-3.93E-18	-6.56E-17	-3.93E-18	-6.56E-17
62	*sa*ta*ha*	0.00001	0.5697	0.5697	-0.5843	-0.5843
63	*sa*ta*ws*	0.009156	-0.2336	-0.2336	0.2396	0.2396
64	*sa*ta*sh*	-0.031	0.5697	0.5697	-0.5843	-0.5843
65	*sa*ta*ch*	0.03959	1.05E-16	1.05E-16	-1.07E-16	-1.07E-16
66	*sa*ta*sd*	0.0384	0.4909	0.4909	-0.5035	-0.5035
67	*sa*ta*cd*	-0.02103	0.2891	0.2891	-0.2965	-0.2965
68	*sa*ha*ws*	0.03982	0.1797	0.1797	-0.1843	-0.1843
69	*sa*ha*sh*	0.01147	-0.4382	-0.4382	0.4494	0.4494
70	*sa*ha*ch*	-0.01552	-8.05E-17	-8.05E-17	8.26E-17	8.26E-17
71	*sa*ha*sd*	0.07496	-0.3776	-0.3776	0.3873	0.3873
72	*sa*ha*cd*	-0.00376	-0.2224	-0.2224	0.2281	0.2281
73	*sa*ws*sh*	0.002782	0.1797	0.1797	-0.1843	-0.1843
74	*sa*ws*ch*	0.01562	3.30E-17	3.30E-17	-3.39E-17	-3.39E-17
75	*sa*ws*sd*	0.03416	0.1548	0.1548	-0.1588	-0.1588
76	*sa*ws*cd*	-0.02964	0.0912	0.0912	-0.0935	-0.0935
77	*sa*sh*sd*	-0.06139	-0.3776	-0.3776	0.3873	0.3873
78	*sa*sh*cd*	-0.148	-0.2224	-0.2224	0.2281	0.2281
79	*sa*ch*sd*	0.03324	-6.94E-17	-6.94E-17	7.11E-17	7.11E-17
80	*sa*ch*cd*	0.03199	-4.08E-17	-4.08E-17	4.19E-17	4.19E-17
	Scale parameter	1.061	NA	NA	NA	NA

[a] NA = Not applicable.

5.7 Conclusion

Due to limitations on the available degrees of freedom for the static, farm-based predictor variables, the EPA developed three alternative EEMs. These three EEMs used the paired combinations of three farm-based variables (i.e., animal type, surface area and farm size) and the continuous variables representing meteorological conditions (i.e., ambient temperature, ambient humidity, wind speed and solar radiation).

After initial evaluation of the three alternatives, the EPA has concluded that additional analysis is needed to develop the lagoon EEMs due to some confounding factors in the available data. Emissions from dairy lagoons during the summer when lagoon emissions are typically higher than the rest of the year are under-represented in the NAEMS. This factor likely causes dairies to appear to have lower emissions than swine. Also, during our evaluation of the three draft EEMs, we concluded that the emissions and surface areas at two of the dairy sites were likely under-represented because all stages of the multi-stage treatment systems at each site were not monitored. This factor likely contributed to the prediction by the *size/sa* EEM that for two farms with the same lagoon surface area, NH_3 emissions are higher for the smaller sized farm. We also observed that the *animal/sa* EEM for both dairy and swine that NH_3 emissions from the farm with 190,000 lbs of animal weight is only 25 percent higher than the farm with 680,000 lbs of animal weight, even though animal weight is 2.8 times as large. Additionally, the reported surface area for swine site IA3A is much smaller than for the other swine sites. The EPA plans to investigate how this disparity in surface area values affects the predictive ability of the EEMs.

The EPA is considering conducting additional analysis to develop the lagoon EEMs that produce emissions more consistent with expectations regarding relationships between the emissions and the predictor variables.

The types of additional analysis under consideration are as follows:

1. Re-examine some suspect data values to determine if they are representative.
2. Consider altering the mathematical form that quantifies interactions between the farm-based variables with meteorological variables.
3. Consider using a weighting scheme whereby some emissions observations are given more weight than others in EEM development. The weighting would account for situations in which the value of a predictor variable for one site falls far from the values for all other sites and has a disproportionate influence on the relationships between that variable and emissions.
4. Consider developing a single static variable that can represent NH3 loading and lagoon surface area. This variable could be constructed using the ratio of nitrogen excretion rates from the Natural Resource Conservation Service (NRCS) to lagoon surface area.

The EPA is seeking recommendations from SAB on the additional proposed analyses.

6.0 REFERENCES

Aneja, V., B. Malik, Tong, Q., Kang, D., and Overton, J., 2001. “Measurement and Modeling of Ammonia Emissions at Waste Treatment Lagoon-Atmospheric Interface,” Journal of Water, Air and Soil Pollution, 1: 177-188.

Aneja, V. P., B. Bunton, J. T. Walker, and B. P. Malik. 2001. Measurement and analysis of atmospheric ammonia emissions from anaerobic lagoons. Atmos. Environ. 35:1949–1958.

EPA (U.S. Environmental Protection Agency). 2001. Emissions from Animal Feeding Operations (Draft). EPA Contract No. 68-D6-0011. Washington, D.C. Available on-line at http://epa.gov/ttn/chief/ap42/ch09/draft/draftanimalfeed.pdf.

Grant, R.H. 2008, Quality Assurance Project Plan for the National Air Emissions Monitoring Study (Open Source Emissions Component), Ver. 4, Purdue Applied Meteorology Laboratory, Department of Agronomy, Purdue University

Harper, L.A. and Sharpe, R.R. 1998. Ammonia emissions from swine waste lagoons in the southeastern U.S. coastal plains. Final Report, Div. of Air Quality, N.C. Dept. of Environment and Natural Resources, Raleigh, NC.

USDA. 1999. Cattle: Final Estimates 1994-1998. Statistical Bulletin 953. U.S. Department of Agriculture (USDA), National Agricultural Statistics Service, Washington, D.C.

USDA 2009, 2007 Census of Agriculture, USDA, National Agricultural Statistics Service. Washington, D.C. available on-line at http://www.agcensus.usda.gov/Publications/2007/Full_Report/usv1.pdf.

USDA’s 2007 Census of Agriculture Dairy Cattle and Milk Production Fact Sheet, Available on-line at http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Fact_Sheets/Production/cattle_and_milk_production.pdf.

USDA’s 2007 Census of Agriculture Hog and Pig Farming Fact Sheet, Available on-line at http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Fact_Sheets/Production/hogsandpigs.pdf

USDA. 2007. Swine 2006, Part I: Reference of Swine Health and Management Practices in the United States, 2006. USDA:APHIS:VS, CEAH. Fort Collins, CO. Available online at http://www.aphis.usda.gov/animal_health/nahms/swine/downloads/swine2006/Swine2006_dr_PartI.pdf

USDA. 2008. Dairy 2007, Part II: Changes in the U.S. Dairy Cattle Industry,1991–2007 USDA-APHIS-VS, CEAH. Fort Collins, CO, Available on-line at http://www.aphis.usda.gov/animal_health/nahms/dairy/downloads/dairy07/Dairy07_dr_PartII.pdf.

National Research Council. 2003. Air Emissions from Animal Feeding Operations: Current Knowledge, Future Needs. National Academy Press, Washington, DC. ISBN: 0-309-08705-8; Available on-line at www.nationalacademies.org.