This is not a peer-reviewed article.
International Symposium on Air Quality and Manure Management for Agriculture
CD-Rom Proceedings of the 13-16 September 2010 Conference
(DoubleTree Hotel, Dallas Texas) Publication date, 13 September 2010
ASABE Publication Number 711P0510cd

GROUND TRUTHING CALPUFF AND AERMOD FOR ODOR DISPERSION FROM SWINE BARNS USING AMBIENT ODOR ASSESSMENT TECHNIQUES

C. G. Henry[1], P. C. D'Abreton[2], R. J. Ormerod[2], G. Galvin[2], S. J. Hoff[3], L. D. Jacobson[4], D. D. Schulte[1] and D. P. Billesbach[1]

ABSTRACT

A collaborative research effort by several institutions investigated the dispersion of odors from a swine production facility. Trained human receptors measured downwind odor concentrations from four tunnel-ventilated swine finishing barns near Story City, Iowa, during twenty measurement events conducted between June and November 2004. Odor concentrations were modeled for short time steps using CALPUFF and AERMOD atmospheric dispersion models to compare predicted and measured odor levels. Source emission measurements and extensive micrometeorological data were collected along with ambient odor measurements using the Nasal Ranger® device (St. Croix Sensory, St. Paul MN), Mask Scentometer, odor intensity ratings, and air sample analysis by dynamic triangular forced-choice olfactometry (DTFCO). AERMOD predictions fit the odor measurements slightly better than CALPUFF with predicted concentrations being about half those predicted by CALPUFF. The Mask Scentometer and Nasal Ranger® measurements related best to the dispersion model output, and scaling factors of 3.0 for CALPUFF and 2.4 for AERMOD suggested for the Nasal Ranger® and 0.5 for the Mask Scentometer (both models). Measurements obtained using the Nasal Ranger®, Mask Scentometer, and odor intensity ratings correlated well to each other, had the strongest linear relationships, and provided slopes (measured: modeled) closest to 1.0. Converting intensity ratings to a dilution to threshold concentration did not correlate and relate as well, and this method was deemed less desirable for ambient odor assessment. Collection of ambient air samples for analysis in a olfactometry laboratory displayed poor correlations with other methods and should not be used to assess ambient odors.

KEYWORDS: CALPUFF, AERMOD, Odor modeling, Nasal Ranger, Mask Scentometer, Odor Intensity Reference Scales, Odor Intensity

INTRODUCTION

Odor issues have become a limiting factor in the viability and growth of livestock and poultry production in the United States. Odor dispersion from livestock facilities is a complicated process that depends on many factors, such as the production system, stocking density, season, localized weather patterns, terrain, and receptor locations relative to the production areas. The National Research Council (NRC, 2003) suggested that one of the two major ways to assess the effects of airborne emissions from animal feeding operations is to replace the current

---

[1] Biological System Engineering, University of Nebraska-Lincoln, 68583, USA

[2] PAE Holmes LTD, Brisbane, Queensland, QLD, 4101, Australia

[3] Department of Agricultural and Biosystems Engineering, Iowa State University, 50011, USA

[4] Department of Bioproducts and Biosystems Engineering, University of Minnesota, 55108, USA

emission factor approach with process-based modeling. Therefore, improved methods and tools are needed to assist in describing the odor risk posed by new and existing facilities on the neighboring community. Such methods and processes would be valuable to the livestock and poultry industry and rural communities when siting new facilities and expanding current production facilities. Moreover, there would also be benefits in the evaluation and adoption of control and mitigation strategies.

Currently, there are several models being used in the United States to evaluate odor dispersion. Researchers at the University of Minnesota used INPUFF-2 (Bee-Line Software Co, Asheville, NC), a US EPA Gaussian puff model described by Peterson and Lavdas (1986), to predict odor levels in the development of the Odor From Feedlots Setback Estimation Tool or OFFSET (Jacobson et al., 2005 and Guo et al., 2005). The University of Nebraska-Lincoln is using AERMOD (AMS/EPA Regulatory Model) to develop an odor risk-assessment tool called the Nebraska Odor Footprint Tool or NOFT (Koppolu et al., 2004; Schulte et al., 2004; Stowell et al., 2005; and Niemeir et al., 2008). AERMOD is a Gaussian dispersion model developed in a joint effort by the American Meteorological Society and the US EPA, as the replacement to ISC3 (Industrial Source Complex 3). Iowa State University has also developed its own model, called CAM (Community Assessment Model) for predicting odor dispersions in a community (Hoff and Bundy, 2003).

The objectives of the work reported in this paper are to 1) develop model-specific scaling factors for CALPUFF and AERMOD associated with measurements made using the following ambient odor assessment techniques: Nasal Ranger®, Mask Scentometer, intensity ratings, and laboratory analysis of collected air samples; and 2) compare the performance of CALPUFF and AERMOD for odor prediction. An underlying goal of this work was to find the model and ambient odor assessment technique combination that gave the best agreement between predicted and observed concentrations.

PREVIOUS WORK

In agriculture, the primary driver behind ground-truthing odor models with field observations has been the development of "simple tools" that can be used to quickly and inexpensively assess the odor risk presented by a proposed facility on neighbors. Such tools are very useful for planning facilities and screening prospective sites, and can be easily used by livestock producers, planning and regulatory officials, and the general public to envision the odor risk of livestock facilities. Historically, such tools were based solely on model results, but more recent work has incorporated observations of trained assessors to calibrate the models by scaling model results.

The US EPA regulatory model, CALPUFF, is a multi-layer, multi-species, non-steady-state puff dispersion model that can simulate the effects of temporal and spatial variability of micrometeorological conditions on pollutant transport, transformation and removal (Scire et al., 2000). In 1991, the American Meteorological Society (AMS) and the US Environmental Protection Agency (US EPA) initiated a formal collaboration with the goal of introducing planetary boundary layer (PBL) concepts into regulatory models. The result of this work is AERMOD, which was intended to replace the Industrial Source Complex Model systems.

In previous analyses of data from an odor monitoring project described in this paper, Modi (2006) and Schulte et al. (2007) compared AERMOD concentration predictions to Nasal Ranger® observations and found overall scaling factors of 1.66, with the model under-predicting the Nasal Ranger® observations. However, these results were only from a subset of ten experiments. This work includes an additional ten experiments (for a total of 20). D'Abreton et al. (2007) modeled downwind odor concentrations for this experiment using the CALPUFF modeling system and found that the 1-minute time step version of CALPUFF was able to mimic the variable nature of odors and that the Nasal Ranger® observations were within the range of model predictions 44% of the time. Henry et al. (2007) analyzed these results further and found a scaling factor of 0.99, implying that a scaling factor may not be needed when CALPUFF predictions are intended to match Nasal Ranger® observations. This paper is a continuation and extension of the work of D'Abreton et al., (2007); Henry et al., (2007); Modi (2006); Schulte et al., (2007).

METHODOLOGY

Trained human receptors measured odor concentrations downwind of four tunnel-ventilated swine finishing barns near Story City, Iowa, during twenty-six measurement events conducted between June and November 2004. A more detailed description of the site characteristics, emission characterization, and meteorological instrumentation is discussed in Henry et al (2007), Schulte et al (2007), Modi (2007), and Henry (2009). Due to micrometeorological equipment failure, data from six of these events were not used, leaving twenty usable events. For each of the twenty 15-minute measurement events, receptors measured odor levels at four locations in the plume, resulting in 80 observations (n=80) for which model predictions and odor measurements could be compared over the same time period and place.

Ambient Odor Assessment Methods

At each of the four downwind measurement locations, ambient odor was assessed using the following techniques:

Nasal Ranger®. Two assessors from Iowa State University trained by St. Croix used the Nasal Ranger® field olfactometer (www.nasalranger.com). Unit setting numbers that corresponded to dilution-to-threshold readings were made twice during each 15-minute assessment period, once at the beginning of the session and the second time at the 7.5-minute mark. The unit settings for the Nasal Ranger® were 0 (no detect), 2, 4, 7, 15, 30, and 60 D/T. The average of each set of four D/T readings made (two assessors x 2 readings) was used in the analysis (n = 80).

Mask Scentometer. During several but not all measurement events, one or two assessors trained by the University of Nebraska used the Mask Scentometer – as described by Sheffield et al. (2004) and Henry (2009). These assessors recorded unit setting numbers that corresponded to dilution-to-threshold readings every 30 seconds during each 15-minute measurement event, for a total of 30 D/T readings per assessor per event. The unit settings used in this work for the Mask Scentometer were 0 (no detect), 0.35, 1, 2, 4.5 and 18 D/T, and the average of the 30 individual D/T readings was used in the analysis. If there was more than one assessor at the receptor location then the arithmetic average of their results was used in the analysis (n = 55).

Odor Intensity Rating Scale (OIRS). Two assessors trained by the University of Minnesota rated odor intensity based on the static-scale method of ASTM Standard E 544-99, "Standard Practices for Referencing Suprathreshold Odor Intensity". A 0-5 scale was used in this experiment based on n-butanol concentrations in air with a geometric progression of three, with 25 ppm representing I=1 and 2,025 ppm representing I=5. This is the same technique used by Jacobson et al., (2000); Jacobson et al., (2003); Nicolai et al., (2000); Zhu et al., (2000). During each 15-minute event, each assessor rated odor intensity 60 times (every 15 seconds). Average intensities for both assessors during the event were averaged and this intensity value was used in the analysis (n=80).

In the data analyses, comparisons were also made using odor concentration values that were predicted from the average intensities. For clarity, these empirically derived odor concentration values are referred to as 'Average intensity-predicted D/T' in this paper. Jacobson et al. (2000) published a relationship between intensity and dilutions to threshold (D/T) as determined from the analysis of odors using a laboratory olfactometer. For swine odors, they used the relationship $D/T_{swine} = 8.367\ e^{1.0781I}$ to obtain D/T values from intensities. Jacobson et al., (2003); Nicolai et al., (2000); Zhu et al., (2000) also used this prediction equation. In the current work, a D/T value was predicted for each measurement event using this equation and the average intensity for the event (n=80).

DTFCO. Ambient air samples were collected in the field for subsequent odor analysis using dynamic triangular forced-choice olfactometry (DTFCO). One air sample was collected in a new unflushed Tedlar 10 L bag during the first four minutes of each 15-minute measurement event. The Iowa State University Odor Lab analyzed odor samples following the ASTM Standard E679-97, “Standard Practice for Determination of Odor and Taste Thresholds by a Forced-Choice Ascending Concentration Series Method of Limits”. The lab was in compliance with the European Standard for olfactometry (CEN, 2003). All samples were analyzed within 24 hours of collection (n=80).

All assessors recorded readings on pre-printed data sheets. Field samples and measurement times were synchronized with the on-site micrometeorological station clock so measurement intervals would correspond to modeled time steps. One challenge with analyzing scentometer data arises from the crisp, nonlinear unit D/T settings. When an odor concentration of 2 D/T is reported by an assessor using a device having unit settings of 2 and 7 D/T, it is very likely that the actual concentration was somewhere between 2 and 7 D/T – so, actual odor concentrations equating to 3 and 6 D/T both would be reported as 2 D/T, which leads to results being skewed downward. To account for this undesired influence, dilution-to-threshold data for the Mask Scentometer and Nasal Ranger® were adjusted using Equation 1 (Sheffield et al., 2004) to obtain geometric average dilution-to-threshold $(D/T)_G$ readings.

Equation 1. $$D/T_{G,n} = 10^{\frac{\log D/T_n + \log D/T_{(n+1)}}{2}}$$

Where n is the device setting number reported by an assessor for a reading, and D/T_n is the D/T specified for that setting (referred to as the ‘unit D/T’). For example, for a reported Nasal Ranger® setting of three (unit D/T of 4 and next higher unit D/T of 7), the $(D/T)_G$ is 5.3. The $(D/T)_G$ used in representing field olfactometer readings are shown in Table 1. Another issue with data from field olfactometers is how to deal with non-detects of odor. The fact that odor was not detected at the lowest device setting does not mean no odor existed in the ambient air, but that it was not perceptible to the assessor. Also, it is not possible to take the geometric average of results that include zeros. For non-detect readings, the $(D/T)_G$ was assumed to be about two-thirds between zero and the first unit D/T on the device (0.2 for the Mask Scentometer and 1.4 for the Nasal Ranger®). No attempt was made to adjust the D/T for the last settings of the devices (18 D/T for the Mask Scentometer and 60 D/T for the Nasal Ranger®) since they are the limits of the instruments (their $(D/T)_G$ could be anywhere between 18/60 and infinity). Using $(D/T)_G$ data increases the odor concentration for an assessment compared to using the unit D/T, which has a justifiable basis. The main drawback to this approach is that when no odor is actually present, a small odor level is reported (i.e. $(D/T)_G$ cannot be less than 0.35 or 1.4). In this situation, it was assumed that an odor was present during the assessments, so there should be little effect of assigning a $(D/T)_G > 0$ for non-detects on the overall results.

Table 1. Geometric Dilutions to Threshold $(D/T)_G$ used for Mask Scentometer and Nasal Ranger®

Mask Scentometer		Setting	Nasal Ranger®	
Unit D/T	Geometric (D/T)	n	Unit D/T	Geometric (D/T)
na	na	7	60	60
18	18	6	30	42.4
4.5	9	5	15	21.2
2	3	4	7	10.2
1	1.4	3	4	5.3
0.35	0.6	2	2	2.8
0 / No detect	0.2	1	0 / No detect	1.4

Modeling Methodology

The AERMOD and CALPUFF models require meteorological data, source emission rate data, facility layout and dimensions and receptor location information. AERMOD is designed to accept hourly micrometeorological data. However, data with a 1-minute time step can be input as hourly data to produce 1-minute predictions, without any model adjustment, when corresponding micrometeorological data are available (Modi, 2006). The meteorological data were averaged every minute and this data was used in the modeling.

Statistical Analysis Methodology

The data from the ambient odor assessments and model predictions were analyzed using bias and error analysis (results not reported in this paper). Initial investigation of the data indicated that it was linear in nature. Additionally, linear regression (R Development Core Team, 2008) was used to develop scaling factors, between the models and respective ambient odor assessment methods. By using both analysis techniques, the best paired ambient odor assessment technique and dispersion model combination can be established. That is the combination with the lowest error and bias, and the combination with the best fitting scaling factor. The slope for regression, or scaling factor, would be a slope nearest 1.0 and that have a coefficient of determination (R_o^2) near 1.0. The coefficient of determination is the proportion of the variability that is accounted by the linear model and describes the goodness of fit of the linear estimated slope. Using this two tiered approach should provide the most reliable technique to use with CALPUFF and AERMOD with the smallest best fitting scaling factor and lowest bias and error.

As a separate analysis, odor methods were compared among themselves using linear regression to develop scaling factors between the different methods. In this analysis, the best fitting slope or scaling factor between methods was sought. The ultimate goal of which was to find the methods that were the most comparable to each other, or equivalent in their assessment of D/T, and if not comparable, how much scaling would be necessary to relate one method to the other (i.e. a Mask Scentometer D/T to a DTFCO lab D/T).

RESULTS AND DISCUSSION

Regression Results

Linear regression analysis results (slopes, coefficients of determination, and standard errors) are shown in Table 2. The slopes represent scaling factors needed to relate values obtained from the various models and odor assessment methods to each other. Traditionally in linear regression analysis, one variable is the independent variable or predictor (x) and a relationship can be found for the dependent variable or response (y). One of the underlying assumptions is that the regressors (x_i) are not contaminated with errors and are independent. In this experiment, this assumption is not valid. So one should base the relationship on the predictor error that is small, to negligible, with respect to the response variable, in order to derive the best relationship possible between methods. Thus, the standard error of the estimate was used as criterion for model selection. The standard error of estimate is a measure of error of prediction. That is the lower the standard error, the higher the precision, and the more preferred model. So each method was regressed as both an independent variable and dependent variable relative to the other methods, as shown in Table 2, and the two regression models were ranked. The model with the lowest error was the better model slope or scaling factor produced from the regression. The slope with a "*" produced the lowest error and is the more precise relationship. Note that the coefficients of determination (R_o^2) are the same for each of the linear models.

From Table 2 one can relate one method to another and assess the scale difference from the different methods. For illustration, the slope between the Mask Scentometer $(D/T)_G$ and Nasal Ranger® $(D/T)_G$ is about one-fifth (0.20), so Nasal Ranger® $(D/T)_G$ readings were about 5 times higher than Mask Scentometer $(D/T)_G$. Meanwhile, the slope of Average intensity-predicted D/T (y) as a function of Nasal Ranger® $(D/T)_G$ (x) was 2.7, however, the Nasal Ranger® $(D/T)_G$ (y) as a function of Average intensity-predicted D/T (x) resulted in a slope of 0.19 and had a lower error (*). Therefore, to relate an assessment made with Average intensity-predicted D/T (y) to an assessment made with a Nasal Ranger® $(D/T)_G$, we would select the stronger relationship, which is 1 Average intensity-predicted D/T = 5.2 Nasal Ranger® $(D/T)_G$ (1/0.19 = 5.2).

Table 2. Regression Results: Slope or Scaling Factor (top), Coefficient of Determination, R_o^2 (middle), and Standard Error (bottom).

(Response) Y► (Predictor) X▼	CALPUFF model D/T	AERMOD model D/T	Nasal Ranger® D/T	Nasal Ranger® (D/T)$_G$	Intensity	Average Intensity-predicted D/T	Mask Scentometer D/T	Mask Scentometer (D/T)$_G$	DTFCO Lab D/T
CALPUFF model D/T		0.80 0.42 0.11	0.99 0.33 0.2	1.19 0.37 0.2	0.09* 0.35 0.01	3.79 0.25 0.07	0.19* 0.47 0.03	0.39* 0.51 0.06	9.06 0.29 1.9
AERMOD model D/T	0.52* 0.42 0.07		0.87 0.38 0.1	1.06 0.44 0.1	0.09* 0.46 0.01	3.32 0.30 0.6	0.14* 0.55 0.02	0.27* 0.52 0.04	5.37 0.39 0.9
Nasal Ranger® D/T	0.33* 0.33 0.05	0.44* 0.38 0.06			0.07* 0.63 0.006	3.30 0.58 0.3	0.12* 0.73 0.01	0.20* 0.49 0.03	4.9 0.43 0.8
Nasal Ranger® (D/T)$_G$	0.31* 0.37 0.05	0.41* 0.44 0.05			0.06* 0.64 0.005	2.73 0.52 0.3	0.11* 0.73 0.01	0.19* 0.55 0.03	4.13 0.45 0.06
Intensity	3.8 0.35 0.60	5.4 0.46 0.7	8.6 0.63 0.8	10.1 0.64 0.9		44.0 0.83 2.3	1.40 0.76 0.1	2.56 0.63 0.3	57.5 0.55 6.9
Average Intensity-predicted D/T	0.07* 0.26 0.10	0.09* 0.30 0.02	0.17* 0.58 0.02	0.19* 0.52 0.02			0.02* 0.64 0.003	0.04* 0.35 0.008	1.28 0.46 0.2
Mask Scentometer D/T	2.5 0.47 0.43	3.87 0.55 0.56	6.1 0.73 0.6	6.6 0.73 0.6	0.54* 0.76 0.05	26.2 0.64 3.1			29.6 0.47 5.9
Mask Scentometer (D/T)$_G$	1.29 0.51 0.21	1.88 0.52 0.29	2.50 0.49 0.4	2.87 0.55 0.4	0.25* 0.63 0.3	9.7 0.35 2.0			15.45 0.47 3.0
DTFCO Lab D/T	0.03* 0.29 0.006	0.07* 0.39 0.01	0.09* 0.43 0.01	0.11* 0.45 0.02	0.009* 0.55 0.001	0.36* 0.46 0.05	0.02* 0.47 0.003	0.03* 0.47 0.006	

* Indicates more precise relationship based on having the lowest standard error. To scale a Nasal Ranger D/T to a Mask Scentometer (D/T)$_G$, multiply the Nasal Ranger value times 0.20 (i.e. 1 Nasal Ranger D/T = 0.20 Mask Scentometer (D/T)$_G$). To scale methods using scaling factors in the light grey boxes, use the inverse slope, for example to relate a Nasal Ranger (D/T)$_G$ to an intensity-predicted D/T, the stronger relationship is 0.19 (rather than 2.7), so multiply the Nasal Ranger value times 5.2 (1/0.19 = 5.2) (i.e. 1 Nasal Ranger (D/T)$_G$ = 5.2 intensity-predicted D/T).

CALPUFF

Shown in Table 2, slopes between CALPUFF and AERMOD results were found to be 0.80 and 0.52, with a slope of 0.52 being the stronger relationship (higher R_o^2). This means that a CALPUFF prediction is about twice that of an AERMOD prediction. The ambient odor assessment method showing the strongest relationship with CALPUFF was the Mask Scentometer ($R_o^2 = 0.47$ for D/T and $R_o^2 = 0.51$ for (D/T)$_G$). The next strongest R_o^2 was for the Nasal Ranger® (D/T)$_G$ ($R_o^2 = 0.37$), Intensity rating ($R_o^2 = 0.35$) and Nasal Ranger® D/T ($R_o^2 = 0.33$). Both Average intensity-predicted D/T and DTFCO had the lowest coefficients of determination ($R_o^2 = 0.25$ and $R_o^2 = 0.29$). The methods with the slopes nearest one, using the lowest standard error, were Nasal Ranger® (D/T)$_G$ (slope=1/0.37=2.7), the Mask Scentometer (D/T)$_G$ (slope=0.39), Nasal Ranger® (slope=1/0.31=3.2), Mask Scentometer D/T (slope=0.19), DTFCO (slope=9.0), Intensity rating (slope=1/0.07=14), and Average intensity-predicted D/T (slope=14.2). These slopes represent the odor assessment method dependent scaling factors for CALPUFF. Based on this dataset and using the R_o^2 and a slope nearest one representing the best relationship as criteria, both the Nasal Ranger and Mask Scentometer appear to be best matched to CALPUFF predictions. From the entire dataset, the Nasal Ranger has a slightly better scaling factor (2.7-3.2), but the Mask Scentometer's scaling factors of 0.19 and 0.39 (slope) are a better fit (R_o^2=0.47-0.51).

It should be noted that the Mask Scentometer data set is much smaller (only half the number of observations) and that the range of the instrument is limited (max assessment of "18 D/T"). The effect of this limitation was studied further, and CALPUFF predictions that were higher than the theoretical limit of 18 Odor Units (assumed equivalent to a D/T) were removed from the dataset (deleted 22 observations with 23 observations remaining). When these high model predictions were removed, which were primarily the closest locations to the barn where the highest odor concentrations were experienced, the slopes improved from 0.19 to 0.28 for Mask Scentometer D/T and 0.39 to 0.52 for the Mask Scentometer $(D/T)_G$ and a modest reduction in the R_o^2 as shown in Table 3. With the high model predictions removed, the Mask Scentometer $(D/T)_G$ has the slope closest to one (0.52) and a coefficient of determination (0.42) better than the Nasal Ranger. While likely not the entire reason, the range limitation of the Mask Scentometer was likely a factor in the results.

Table 3. Slope (top), Coefficient of determination R_o^2, and standard error for CALPUFF predictions less than 18 OU.

	Mask Scentometer D/T	Mask Scentometer $(D/T)_G$
CALPUFF	0.28* 0.47 0.05	0.52* 0.42 0.1
AERMOD	0.22* 0.44 0.04	0.54* 0.56 0.08

** Indicates the lowest standard error, where Mask Scentometer measurements are y, and model predictions are the x variables.*

When $(D/T)_G$ is used instead of D/T for the Nasal Ranger® and Mask Scentometer, the slopes for the Nasal Ranger® modestly decreased from 0.33 to 0.31 yet for the Mask Scentometer improved from 0.19 to 0.39 (or 0.28 to 0.52 when high model predictions were removed). In both cases R_o^2 improved slightly. This showed that $(D/T)_G$ provides a better result for the Mask Scentometer, but is not necessarily helpful in improving results for the Nasal Ranger®. In summary, the ambient odor assessment methods showing the best relationship to the CALPUFF model predictions were the Nasal Ranger® and Mask Scentometer. Curran et al., (2007) found the average predicted to measured mean concentration ratio on the sampling days to vary from 1.4 to 9.37 (which would relate to model scaling factors of 0.7 and 0.10), for an OIRS that included a 1-butanol scale and using VDI 3940 as a guideline when comparing ISC3 and CALPUFF for odors. Li and Guo (2006) found scaling factors from 1.2-7.9 using CALPUFF and an OIRS (0-8 scale). Zhu et al., (2001) found scaling factors of 10 for manure storages and 35 for buildings for INPUFF-2 using the same OIRS (0-5 scale) as this study. In this study, the most comparable method, Average intensity-predicted D/T, produced a scaling factor of 14 for CALPUFF.

AERMOD

The method with the strongest R_o^2 is the Mask Scentometer (0.55 for D/T and 0.52 for $(D/T)_G$). The next best R_o^2 was Intensity Rating (0.46), the Nasal Ranger® $(D/T)_G$ (0.44), DTFCO lab D/T (0.39), Nasal Ranger® D/T (0.38), and Average-intensity-predicted D/T (0.30). The methods with the slopes nearest one, using the more precise relationships(standard error), were Nasal Ranger® D/T (1/0.44=2.3), Nasal Ranger® $(D/T)_G$ (1/0.41=2.4), the Mask Scentometer $(D/T)_G$ (0.27), DTFCO (5.3), Mask Scentometer D/T (0.14), and Intensity rating and Average-intensity-predicted D/T both the same (1/0.09=11.1). These slopes represent the odor assessment method dependent scaling factors for AERMOD. Based on this dataset and using the R_o^2 and a slope nearest one representing the best relationship as criteria, both the Nasal Ranger and Mask Scentometer appear to be best matched to AERMOD predictions because they have the strongest R_o^2 and slopes closest to one. From the entire dataset, the Nasal Ranger has a slightly better scaling factor (2.3-2.4), but the Mask Scentometer's scaling factor (slope) is a better fit (R_o^2=0.52-0.55). Again, like the CALPUFF results, $(D/T)_G$ related better to the AERMOD predictions for the Mask Scentometer, but does not appear to have any substantial effect for the Nasal Ranger®. Clearly, the intensity rating, DTFCO Lab D/T and the Average-intensity-predicted D/T do not relate as well with AERMOD predictions when using R_o^2 and nearness to a 1:1 slope as criteria.

Similar to CALPUFF, when the high AERMOD predictions greater than 18 OU (assumed equivalent to a D/T), were removed from the dataset, the slopes improved as shown in Table 3. The more precise model as determined by the lowest standard error was by using the model prediction as the independent variable and the Mask Scentometer measurement as the dependent variable. A modest reduction in the R_o^2 and a slope improvement from 0.14 to 0.22 (Mask Scentometer D/T) and from 0.27 to 0.54 (Mask Scentometer $(D/T)_G$ using the dataset with the high model predictions removed for the Mask Scentometer D/T and $(D/T)_G$ respectively.

Suggested CALPUFF and AERMOD Model Scaling Factors

Table 4 shows the suggested scaling factors to be applied to CALPUFF and AERMOD odor predictions for livestock building sources found from this study. These scaling factors would be applicable to modeling predictions that were done at one-minute time steps, additional scaling factors may be needed for hourly model predictions.

Table 4. Suggested Scaling Factors for CALPUFF and AERMOD

Model	Nasal Ranger® (either D/T or $(D/T)_G$)	Mask Scentometer D/T	Mask Scentometer $(D/T)_G$	Average-intensity-predicted D/T
CALPUFF	3.0	0.39	0.52	14
AERMOD	2.4	0.22	0.54	11

Odor Methods Results

The Spearman's Rank Correlation coefficients (ρ) were used to indicate the strength and direction of the linear relationship between two random variables. The Spearman's Rank Correlation is a special case of the Pearson product-moment coefficient, in which the two sets of data are ranked before calculating the coefficient. In Table 5 the raw scores are converted to ranks and the differences between the ranks of each observation on the two variables were calculated using the "cor.test" command in the R statistical package (R Development team, 2008). The correlation coefficient lies between -1 and 1, with 1 indicating a strong linear relationship (-1 strong inverse relationship, i.e. negative slope) and a 0 indicating no linear relationship. The Spearman's Rank correlation was used because non-parametric tests are considered to be more robust because they do not rely on the assumption that the data comes from a normal distribution of odor levels (i.e. range of possible values and the probability that the measurement was in that range). Also since the intensity rating can be considered an ordinate scale, the non-parametric test (Spearman's ρ) is more applicable and allows for an all methods comparison. The null hypothesis is that no correlation exists between odor measurement methods, the alternative is that correlation (a relationship) exists between odor measurement methods (P values must be less than α=0.05 and α=0.10 to indicate significant correlation exists) and that the relationship is linear.

Table 5. Spearman's Correlation Coefficient, ρ

	Intensity Rating (0-5)	Average-intensity-predicted D/T	Mask D/T	Mask $(D/T)_G$	DTFCO Lab D/T
Nasal Ranger® D/T	0.50*	0.50*	0.53*	0.32*	0.15
Nasal Ranger® $(D/T)_G$	0.48*	0.48*	0.52*	0.33*	0.12
Intensity Rating (0-5)		1.0*	0.53*	0.26**	0.09
Average-intensity-predicted D/T			0.53*	0.26**	0.09
Mask D/T					0.01
Mask $(D/T)_G$					0.09

** Denotes P<α=0.05, there is a correlation between methods. ** P<α=0.10*

When correlating the odor assessment methods in Table 5 some patterns emerge. First it is apparent that DTFCO Lab D/T does not correlate to any of the other odor measurement methods. The strongest correlations exist between the Mask Scentometer D/T (0.53), Intensity Rating (0.53), Average-intensity-predicted D/T (0.53) and both Nasal Ranger® D/T (0.53) and $(D/T)_G$ (0.52). As expected intensity rating and Average intensity-predicted D/T are perfectly correlated

(1.0). The Nasal Ranger® also correlates well with intensity rating (0.48 and 0.50), and Average intensity-rating D/T (0.48 and 0.50). The Mask Scentometer $(D/T)_G$ is not as strongly correlated as Mask Scentometer D/T, but is still significantly correlated to the Nasal Ranger® and to intensity rating and Average intensity-predicted D/T (α= 0.10).

To relate the methods to each other, linear regression was performed as shown previously Table 2. The method with the strongest fit are intensity rating (0-5) and Average intensity-predicted D/T (R_o^2=0.83). The next best R_o^2 is between intensity and the Mask Scentometer D/T (R_o^2=0.76). Another good R_o^2 is 0.73 between both Nasal Ranger® D/T and $(D/T)_G$ and Mask Scentometer D/T, although correlation degrades when compared to Mask Scentometer $(D/T)_G$ (R_o^2=0.54 and 0.55). The next best correlations are between intensity and the Nasal Ranger® $(D/T)_G$ (R_o^2=0.64) and the Mask Scentometer D/T (0.76) and $(D/T)_G$ (0.63), and DTFCO (0.55). In general DTFCO Lab D/T has the weakest R_o^2 to any of the other methods.

The slope for regression of two perfectly comparable methods - methods that both produce the same result - would be 1.0 and methods that have a coefficient of determination near 1.0. The methods with the slope nearest one are, Mask Scentometer D/T and intensity (1/0.54=1.9), DTFCO lab D/T and Average intensity-predicted D/T ((1/0.36=2.8), Mask Scentometer $(D/T)_G$ and intensity (1/0.25=4), Mask Scentometer $(D/T)_G$ to Nasal Ranger® D/T (0.20) and $(D/T)_G$ (0.19), Nasal Ranger $(D/T)_G$ to Average intensity-predicted D/T (1/0.19=5.2) and Mask Scentometer D/T to Nasal Ranger® D/T (0.12) and $(D/T)_G$ (0.11).

In the CALPUFF and AERMOD analysis, the limitation of the Mask Scentometer was evaluated by removing data points that were greater than the last setting of the device, 18 D/T. Data points were removed when the Nasal Ranger® assessments were greater than 18 for both D/T and $(D/T)_G$ (n=31). This resulted in the slopes nearly doubling, from 0.12 to 0.20 for D/T and 0.20 to 0.48 for $(D/T)_G$. The R_o^2 is still strong at 0.63 and 0.60 for D/T and $(D/T)_G$. We conclude from this analysis that the limited range of the Mask Scentometer is a factor in the results. Additionally, it seems logical that the Mask Scentometer would "average" out a few high D/T values, where just one high or low D/T from the Nasal Ranger® could skew the results (only two assessments per session were taken). Also, there less data available from the Mask Scentometer readings (n=55) than for the intensity and Nasal Ranger® assessments (n=80), so with more replication, the results may be different.

Table 6. Device Limitation: Regression Results, slope (top), Coefficient of Determination, R_o^2 (middle) and Standard Error (bottom) for Nasal Ranger® assessments less than 18 D/T.

	Mask Scentometer D/T	Mask Scentometer $(D/T)_G$
Nasal Ranger D/T	0.20 0.63 0.03	
Nasal Ranger $(D/T)_G$		0.48 0.60 0.08

SUMMARY AND CONCLUSIONS

Odor emissions from four swine barns were intensively sampled during a series of twenty 15-minute experiments, where ambient downwind odors were assessed using the Nasal Ranger®, Mask Scentometer, intensity, and DTFCO. Micrometeorological parameters were measured with a cup anemometer and wind vane as well as a high frequency sonic anemometer, temperature and relative humidity at four elevations, short-wave net solar radiation and soil heat flux. Predictions from measured odor emission rates and micrometeorological conditions were generated with CALPUFF and AERMOD and compared to one another and to the ambient odor assessment techniques. The following conclusions resulted:

1. It is clear from this work that results of ambient odor assessments depend on the method and device employed and, while results for one method may correlate and relate well to another method, they are not the same, even if the results are both reported as dilutions to threshold (D/T). When D/T are reported or utilized (i.e. modeling, regulations) they should

be reported as a D/T evaluated by the method used (i.e. a 7 D/T was assessed using a Mask Scentometer). In light of this research, it is clear that standards for ambient odor assessment are needed.

2. In general, the ambient odor assessment methods showed a slightly better relationship to AERMOD than CALPUFF. CALPUFF predictions were about twice that of AERMOD predictions. Scaling factors for the three ambient odor assessment methods that performed the best (lowest bias and error, regression slope nearest one, and coefficient of determinations) were The Nasal Ranger® and Mask Scentometer. Intensity rating was also good to a lesser extent, but is less practical because it is not a comparable measurement to an Odor Unit. The recommended scaling factors for CALPUFF and AERMOD are 3.0 and 2.4 for The Nasal Ranger® D/T and $(D/T)_G$ assessments. For the Mask Scentometer CALPUFF scaling factors were found to be 0.52 for $(D/T)_G$ and 0.39 for D/T and AERMOD scaling factors were found to be 0.54 for $D/T)_G$ and 0.22 for D/T. Scaling factors for Average intensity-predicted D/T were found to be 14 for CALPUFF and 11 for AERMOD. Power terms, n for peak to mean ratios, were found to be within the range of those reported in the literature.
3. The Spearman's Correlation Coefficient (ρ), linear regression, and standard error were used to develop and find the best fitting relationships between the odor methods used in this study between ambient odor methods. In general an Average intensity-predicted D/T is about a three to five times that of a Nasal Ranger® Assessment and a Mask Scentometer $(D/T)_G$ is about a fifth to a half of a Nasal Ranger® Assessment.
4. The range of the Mask Scentometer appears to be a factor in the results, and is a likely reason for the scaling factors found in this study to be less than 1.0. When high values were removed from the dispersion model and Nasal Ranger® data sets the scaling factors to the Mask Scentometer improved (moved close to 1.0) when compared to the entire dataset.
5. Geometric average $(D/T)_G$ of the Mask Scentometer and Nasal Ranger® setting was used in this work. This relationship had a pronounced effect on the Mask Scentometer results, but not on the Nasal Ranger Results. This is likely due to the difference in the number of observations taken during the averaging periods. Field olfactometers (specifically the Mask Scentometer, but this may apply also to the Nasal Ranger) that take repeated measures over a period of time, do appear to benefit by having better agreement with dispersion models and other odor methods, from the application of $(D/T)_G$ to their reported scales.
6. Our recommendation is to use the Nasal Ranger® or Mask Scentometer when ground-truthing AERMOD and CALPUFF odor concentration predictions. It provides an overall result that is very near model predictions, when an adequate amount of both ambient and emission data are collected. OIRS and DTFCO had larger scaling factors and weaker relationships to models than the Nasal Ranger® and Mask Scentometer. In general, AERMOD has a slightly better agreement over CALPUFF, but not likely significant, so we cannot recommend one model over the other. While we ran the models at one-minute intervals as opposed to a longer averaging period, we did not evaluate if this was a factor in our results. This work has the most implication to those that model odor dispersion with CALPUFF and AERMOD and use ambient odor techniques to verify model results.

ACKNOWLEDGEMENTS

The project was funded by a USDA CSREES NRI Integrated Programs grant, Verification of Odor Dispersion Modeling for Siting Livestock and Poultry Production Systems, and received additional support from the Agricultural Research Divisions of the respective universities.

REFERENCES

1. American Society for Testing and Materials (ASTM). 1991. E679-97. Standard practices for determination of odor and taste thresholds by a forced-choice ascending concentration series method of limits. *Annual Book of ASTM Standards.* Philadelphia, PA.: American Society of Testing and Materials.
2. American Society for Testing and Materials (ASTM). 1999. E 544-99. Standard practices for referencing suprathreshold odor intensity. American Society for Testing Materials, Philadelphia, PA.
3. Committee for European Normalization (CEN) 2003. CEN/TC264/WG2/N222/e Air Quality – Determination of Odour Concentration by Dynamic Olfactometry, Brussels, Belgium. European Committee for Standardization.
4. Curran, T.P., V. A. Dodd, and W. L. Magette. 2007. Evaluation of ISC3 and CALPUFF Atmospheric Dispersion Models for Odor Nuisance Prediction. Paper number 074181. Annual Internation Meeting, Minneapolis, MN, 17-20 June. ASAE, St. Joseph, MI.
5. D'Abreton, P. C., C Henry and R Ormerod. 2007. Downwind Odour Predictions from Frou Swine Finshing barns using CALPUFF. Consultancy report for Iowa State University. Prepared by Pacific Air and Environment Pty, Ltd., South Brisbane/Sydney, Australia.
6. Henry, C.G. 2009. Development of the Mask Scentometer, a Comparison of Ambient Odor Assessment Methods, and their Application in Ground Truthing Atmospheric Dispersion Models. A dissertation. University of Nebraska at Lincoln, Lincoln, Nebraska, USA. This document can be accessed at: http://digitalcommons.unl.edu/biosysengdiss/6/
7. Henry, C. G., P C. D'Abreton, R. J. Ormerod, G. G. Galvin, S. J. Hoff, L. D. Jacobsen, and D.D. Schulte. 2007. Downwind Odor Predictions From Four Swine Finishing Barns Using CALPUFF. In Proceedings of the International Symposium on Air Quality and Waste Management for Agriculture. September 16-19, 2007 Omni Interlock Resort Broomfield, Colorado.
8. Hoff, S. J. and D. S. Bundy 2003. Modeling odor dispersion from multiple sources to multiple receptors. In: Proceedings of the International Symposium on Gaseous and Odour Emissions from Animal Production Facilities. Norsens, Denmark. June 1-4, 2003. pp 331-339.
9. Jacobson, L.D., H. Guo, D.R. Schmidt, R.E. Nicolai, J. Zhu and K.A. Janni. 2000. Development of an odor rating system to estimate setback distances from animal feedlots: Odor from Feedlots-Setback Estimation Tool (OFFSET). Paper No. 00-4044. Milwaukee, WI, ASAE.
10. Jacobsen, L. D., H. Guo, D. R. Schmidt, R. E. Nicolai, and J. Zhu 2003. Development of OFFSET (Odor From Feedlots Setback Estimation Tool) for determination of odor annoyance free setback distance from animal production sites. In: Proceedings of the International Symposium on Gaseous and Odour Emissions from Animal Production Facilities. Horsens, Denmark, June 1-4, 2003. pp 320-330.
11. Koppolu, L., D. Schmidt, D. D. Schulte, and L. Jacobson 2004. Development of scaling factors (peak-to-mean ratios) through dispersion modeling with AERMOD and field-based odor measurements for livestock facilities. ASAE paper # 044196. ASAE, St. Joseph, MI.
12. Li, J. and Guo H. 2006. Evaluating the Impact of Computation Time Interval on Livestock Odour Dispersion using a Computation Fluid Dynamics Model. Written for presentation at the CSBE/SCGAB 2006 Annual Conference, Edmonton Alberta, July 16-19. Paper number 06-173, The Canadian Society for Bioengineering.
13. Modi, M. 2006. Modeling odor dispersion form a swine facility using AERMOD. Unpublished MS thesis. Department of Biological Systems Engineering. University of Nebraska Lincoln, Lincoln, NE. 86p

14. Niemeir, K, R., C. A. Powers, R.R. Stowell, D.D. Schulte. 2008. Validating the Odor Footprint Tool using Field Data. Written for presentation at the 2008 ASABE Annual Meeting, Providence, Rhode Island, June 29-July 2, 2008. Paper number 084577.
15. Nicolai, R. E., C. J. Clanton, and H. Guo. 2000. Modeling the relationship between threshold and intensity of swine odors. *Second International Conference on Air Pollution from Agricultural Operations.* St. Joseph, Mich.: ASAE.
16. National Research Council. 2003. The Scientific basis for estimating emissions from animal feeding operations. National Research Council. Washington, D. C.
17. Peterson, W. B. and L. A. Lavas 1986. INPUFF-2.0. A multiple source Gaussian puff dispersion algorithm. Users Guide. EPA/600/8-86/024. August.
18. R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
19. Scire, J.S., Strimaitis, D.G. and Yamartino, R.J. (2000) A user's guide for the CALPUFF dispersion model (version 5), Earth Tech, Inc., Concord, MA.
20. Schulte, D. D., L. Koppolu, R. K. Koelsch, and D. P. Billesbach. 2004. Odor Footprint Modeling for Confinement Animal Feeding Operations. Water Environment Federation. Presented at the Annual Wastewater Management Association, Bellevue, WA. April 17th, 2004.
21. Schulte, D.D., M.R. Modi, C.G. Henry, R.R. Stowell, D.P. Billesbach, S.J. Hoff and L.D. Jacobson. 2007. Modeling Odor Dispersion from a Swine Facility using AERMOD. In Proceedings of the International Symposium on Air Quality and Waste Management for Agriculture. September 16-19, 2007 Omni Interlock Resort Broomfield, Colorado.
22. D. Parker 2004. Evaluation of Field-based Odor Assessment Methods. Water Environment Federation and the Air and Waste Management Association Conference on Odor and Air Quality. March 19-21, 2004. Tacoma, WA.
23. Stowell, R R. L. Koppolu, D.D. Schulte and R. K. Koelsch. 2005. Applications of Using the Odor Footprint Tool. Presented at the VII International Livestock Environment Symposium, Beijing, China, May 18-20, 2005. ASAE Pub #701P0205.
24. Zhu, J., L. D. Jacobson, D. R. Schmidt, and R. Nicolai. 2000. Evaluation of INPUFF-2 model for predicting downwind odours from animal production facilities. Applied Engineering in Agriculture. 16(2): 159-164.