Comparison of Ambient Odor Assessment Techniques in a Controlled Environment

C. G. Henry, D. D. Schulte, S. J. Hoff, L. D. Jacobson, A. M. Parkhurst

Abstract. *This article compares results of using dynamic triangular forced-choice olfactometry (DTFCO), the Mask Scentometer, the Nasal Ranger, and an odor intensity reference scale (OIRS) to assess odors in a controlled-environment chamber in the Iowa State University Air Dispersion Laboratory. The methods were used to assess 13 odor levels in the chamber. Swine manure mixed with water was used to vary the odor levels. DTFCO did not correlate well to the other ambient odor assessment methods. Predicting dilution to threshold (D/T) using intensity ratings compared to using intensity ratings directly degraded the coefficient of determination (R_o^2) through zero with the other methods in all cases. Average intensity-predicted D/T, the Mask Scentometer, and the Nasal Ranger correlated well with each other, with strong R_o^2 values (greater than 0.85) and regression slopes near 1, and the session means were not found to be significantly different (α = 0.05). Using the geometric means of the device D/T settings, $(D/T)_G$, improved the R_o^2 values between the other methods and the Nasal Ranger and Mask Scentometer. Average intensity-predicted D/T values were three to four times higher than Nasal Ranger assessment ($(D/T)_G$ and D/T, respectively), and Nasal Ranger $(D/T)_G$ was roughly five times higher than Mask Scentometer $(D/T)_G$.*

Keywords. *D/T, DTFCO, Field olfactometry, Intensity, Mask Scentometer, Nasal Ranger, OIRS.*

Primary difficulties with assessing ambient odors are the low concentrations of odor commonly experienced and the rapidly fluctuating conditions that occur over time. Laboratory-based dynamic triangular forced-choice olfactometry (DTFCO) has generally been the accepted standard method for measuring odor concentrations. In the ambient atmosphere, though, odor concentrations are very low, and DTFCO using Tedlar bags typically is more effective at assessing odors at higher concentrations (20 to 60 D/T or greater) than at the low concentrations encountered downwind from an odor source (Parker et al., 2003). Koziel et al. (2004) and Trabue et al. (2008) both reported losses of odor compounds in Tedlar bags used for olfactometry analysis. Additionally, analyzing air samples with DTFCO can be very expensive. Field olfactometers and odor intensity ratings have the advantage of being less expensive methods for obtaining field data over a longer period of time, making them attractive in calibrating and verifying models, as well as making general assessments of odor (Sheffield and Ndegwa, 2008). In some instances, field olfactometry may be used in conjunction with laboratory-based methods. For example, air samples from an odor source may be collected and analyzed in an olfactometry laboratory to quantify source emissions rates, while field olfactometry is used to assess odor transport in the surrounding area.

Field olfactometers available for use today include the Box Scentometer manufactured by the Barnebey Sutcliffe Corporation (purchased in 2004 by Calgon Carbon Corp., Pittsburgh, Pa.; www.calgoncarbon.com), the Nasal Ranger manufactured by St. Croix Sensory (Lake Elmo, Minn.; www.nasalranger.com), and the Mask Scentometer, also referred to as a facial field olfactometer, an instrument developed by Sheffield et al. (2004a, 2004b) and improved by Henry (2004, 2009). Finally, intensity ratings based on an odor intensity reference scale (OIRS) may be used as predictors of odor concentration (ASTM, 1999a).

Previous Work

Sheffield et al. (2004a, 2004b) investigated differences between the Mask Scentometer, Nasal Ranger, Box Scentometer, in-field intensity, and in-lab intensity (from Tedlar bags) field assessment techniques with DTFCO at five agricultural and industrial sources using a group of eight assessors to make measurements. Their study evaluated the variability of responses of the devices and methods and found that the Nasal Ranger and laboratory-based olfactometry exhibited the least amount of variability across the odor sources. Sheffield et al. (2007) performed odor assessments at 38 dairies and 15 feedlots in Idaho. They assessed odors using the Nasal Ranger and intensity ratings with n-butanol as the reference odorant. They found a moderate correlation between dilution to threshold (D/T) and H_2S/total reduced sulfur (TRS), which appeared to increase slightly with receptor distance from the

Submitted for review in June 2010 as manuscript number SE 8612; approved for publication by the Structures & Environment Division of ASABE in July 2011. Presented at the 2010 International Symposium on Air Quality and Manure Management for Agriculture (ASABE Publication No. 711P0510cd).

The authors are **Christopher G. Henry, ASABE Member,** Extension Engineer, and **Dennis D. Schulte, ASABE Member,** Professor, Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska; **Steve J. Hoff, ASABE Fellow,** Professor, Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, Iowa; **Larry D. Jacobsen, ASABE Fellow,** Professor, Extension Engineer, Department of Bioproducts and Biosystems Engineering, University of Minnesota, Minneapolis, Minnesota; and **Ann M. Parkhurst,** Professor, Department of Statistics, University of Nebraska-Lincoln, Lincoln, Nebraska. **Corresponding author:** Christopher G. Henry, 217 Chase Hall, University of Nebraska, Lincoln, NE 68583-0726; phone: 402-472-6529; e-mail: chenry1@unl.edu.

  ISSN 2151-0032

source. McGinley and McGinley (2003) compared the Barnebey Sutcliffe Box Scentometer and the Nasal Ranger in an environmentally controlled room. A hydrogen sulfide generator was used to vary the odor levels, while three Nasal Ranger assessors and one Box Scentometer user evaluated the odor in the room. They found high correlation (r = 0.82, *n* not reported) between the Box Scentometer and the Nasal Ranger, and no significant difference was found between assessors (p = 0.309). The field olfactometers yielded hydrogen sulfide thresholds of 0.5 to 2.0 ppb. Laboratory olfactometry (DTFCO) yielded comparable thresholds of 0.45 to 0.9 ppb, and the McGinley's deemed their results consistent with other published values.

Newby and McGinley (2004) compared Nasal Ranger, Barnebey Sutcliffe Box Scentometer, and laboratory-based olfactometry for assessing odor in the field. They found no significant difference between the Box Scentometer and a pre-production Nasal Ranger at a 95% confidence interval (p = 0.06) and a Pearson correlation coefficient of 0.82. They found that the Missouri regulatory limit of 110 D/T (their actual mean was 106.5 D/T) using laboratory olfactometry equated to 7 D/T observed with a Box Scentometer. According to the state statute, a 7:1 D/T observed with a scentometer is a trigger for an olfactometry sample (DTFCO) to be taken. The purpose of their work was to show that Box Scentometer readings and D/T from olfactometry analysis of samples were not comparable (i.e., a different standard was needed for olfactometry analysis). Brandt et al. (2008) used laboratory olfactometry and field assessment with a panel using Nasal Ranger field olfactometers to assess odor level differences between manure land application methods. They found that the laboratory olfactometry results were about 2.5 times higher than the assessments by the Nasal Ranger panel. Bokowa (2008) compared the Nasal Ranger with laboratory olfactometry analysis and found that the Nasal Ranger gave odor detection threshold values that were two to four times lower than laboratory olfactometry. Henry et al. (2005) and Henry et al. (2006) compared back-to-back field assessments taken downwind of swine lagoons and beef feedlots with a Mask Scentometer followed by an OIRS. The best fitting relationship between the methods was a power relationship of intensity D/T = 23.6 (Mask D/T)$^{0.51}$, with an R^2 value of 0.52. Essentially an order of magnitude difference was found between methods in these studies.

One problem relating to the results of these initial studies using the Mask Scentometer is that the dilution ratios were assumed to be similar to those of the Box Scentometer. Henry et al. (2011) reported that the dilution ratios were actually different. It should also be noted that the calculation of D/T based on the intensity ratings from the Henry (2004) and Henry et al. (2005) studies is consistent with the term "intensity-predicted D/T" in this article.

Purpose of Work

In spite of the efforts reported above, the measurement of ambient odors is an imprecise science and has limitations. One of the challenges of ambient odor assessment is that there is no standard method to relate one odor assessment technique to another. That is, the reported dilution to threshold of one instrument or method is not currently comparable with that of another. Much odor work has been done with a plethora of methods, yet it is currently not possible to determine how or if the results from these various methods can be related. The objectives of this experiment were to compare four ambient odor assessment techniques (DTFCO, Nasal Ranger, Mask Scentometer, and OIRS) under controlled conditions, and identify relationships between the data produced using these methods.

Materials and Methods

A series of 13 odor assessment sessions were conducted in a controlled laboratory environment at the Iowa State University Air Dispersion Laboratory in May and June 2004. The number of assessments performed for each method was based on the amount of time needed to perform as many odor assessments as could be reasonably performed in a 10 min period. In each session, the following assessment methods were used:

DTFCO

Dynamic triangular forced-choice olfactometry (DTFCO) was used to analyze air samples collected in the chamber in new, triple-flushed with zero air, unbaked Tedlar bags (10 L) during the first 4 min of each 10 min assessment session. Sampling and analysis followed ASTM Standard E679-99 (ASTM, 1999b). The Iowa State University odor laboratory analyzed the air samples using DTFCO. All samples were analyzed to determine a panel D/T within 24 h. The lab was in compliance with the European Standard for olfactometry (CEN, 2003). A single olfactometry sample was used for each session to characterize the odor level for DTFCO.

Nasal Ranger

Assessors from Iowa State University (ISU) were trained by St. Croix Sensory to use the Nasal Ranger field olfactometer. Odor assessments were made twice during each 10 min assessment session: once shortly after entering the room and again 5 min after entering the room. None of the ISU panelists who assessed odors with the Nasal Ranger participated as panelists for the olfactometry analysis (DTFCO).

Mask Scentometer

Assessors trained by the University of Nebraska used the Mask Scentometer field olfactometer developed by Sheffield et al. (2004a, 2004b) and Henry (2004) to assess odors every 30 s during each 10 min session. In the analysis of data, D/T settings were assigned as specified by Henry (2009).

Intensity Rating (OIRS)

Assessors were trained by the University of Minnesota to rate odor intensity using an odor intensity reference scale (OIRS) based on the static scale method of ASTM Standard E 544-99 (ASTM, 1999a). A scale of 0 to 5 was used in this experiment based on n-butanol in air concentrations, using 25 ppm to represent I = 1, 75 ppm for I = 2, 225 ppm for I = 3,675 ppm for I = 4, and 2025 ppm for I = 5. Assessors could use half steps (i.e. 1.5) if they felt the odor intensity was between 1 and 2). Assessments were taken every 15 s, which resulted in 40 assessments taken during each experiment. Field intensity data were analyzed as raw data (intensity) and converted to D/T using two techniques described later and referred to as intensity-predicted D/T and average intensity-predicted D/T.

For the Nasal Ranger, Mask Scentometer, and OIRS methods, three to five individuals were randomly spaced within a (6.8 m × 6.8 m, 20 ft × 20 ft) room located at the Iowa State University Air Dispersion Laboratory. Swine manure from a deep pit finisher was used as the odor source. The manure was diluted with water to achieve the range of odor levels that the researchers felt would be experienced in the field. Air was drawn through an inlet and across the room by exhaust ventilation fans located at the back of the room. A plenum was installed to create uniform airflow across the room. Odor levels were presented in random order for each session. All assessors began their assessments at the same time (a lead assessor began and stopped all assessors).

The experiment was conducted over a period of two days with six 10 min odor sessions conducted the first day and seven sessions on the second day (13 total). On the first day (first six sessions), three assessors used Mask Scentometers, three assessors used Nasal Rangers, and five assessors rated odor intensity. On the second day (last seven sessions), five assessors used Mask Scentometers, five assessors used Nasal Rangers, and four assessors rated odor intensity.

A relationship first used by Sheffield et al. (2004a, 2004b) was used to obtain a geometric average dilution to threshold $(D/T)_G$ for the field olfactometers (Mask Scentometer and Nasal Ranger). The results are shown in table 1. This was done to normalize the peaks and keep extremely high or low values from skewing the results.

Intensity data were used to predict D/T, and the resulting "intensity-predicted D/T" values were used to compare methods. Jacobson et al. (2000) published a relationship between intensity and D/T determined from the analysis of odor concentration using a laboratory olfactometer. For swine odors, they used the following relationship to predict D/T as a function of odor intensity (i):

$$D/T_{swine} = 8.367e^{1.0781i} \quad (1)$$

This relationship was applied to the intensity rating data in two ways. The first way used the equation to predict a D/T value for each individual assessor observation (reported intensity value). The average D/T for each assessor's series of observations was then used for the session to determine an average predicted D/T, which is referred to as "intensity-predicted D/T".

The second way took the average of the intensity rating values and then used the same equation applied to each individual assessor's average intensity ratings (0 to 5) for the session to predict an "average intensity-predicted D/T." This is the same technique used by Jacobson et al. (2000), Jacobson et al. (2003), Nicolai et al. (2000), and Zhu et al. (2000).

Table 1. Geometric dilutions to threshold, $(D/T)_G$, used for the Mask Scentometer and Nasal Ranger.

Mask Scentometer D/T		Setting	Nasal Ranger D/T	
Unit	Geometric		Unit	Geometric
--	--	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/non-detect	0.2	1	0/non-detect	1.4

Results and Discussion

A detailed statistical analysis was completed. Raw data were checked using lack-of fit in SAS (SAS, 2008) for linearity. To screen for bias, a test for interaction between days and sessions was checked, and an analysis of variance (ANOVA) was used to test for variation between methods. A few individual assessors were deemed to have bias and were removed from the dataset. Using the R statistical package (R Development Core Team, 2008), Pearson's product-moment correlation coefficient and Spearman's rank correlation (ϱ) were used to indicate strength and direction of the linear relationship, and linear regression (coefficient of determination, forced intercept through zero, R_o^2) was performed to develop a relationship between methods. Developing statistical relationships between ambient odor assessment methods was complicated by the fact that a different number of odor assessors used each method. Because of the different number of observations available for each method across the sessions, only the session means for each method were used in the statistical analyses.

From the results shown in table 2, several general trends emerge. Most notably, none of the data obtained using field methods correlated well with DTFCO Lab D/T. We theorize that the primary difference between DTFCO and the field methods is that the background odor levels in the collection bags interfere with the low odor concentration observed in the ambient atmosphere. Good correlations existed, as expected, between the intensity ratings and intensity-predicted D/T and average intensity-predicted D/T. Good correlations were found between intensity ratings and Mask Scentometer $(D/T)_G$ (0.84 to 0.86) and between intensity ratings and Nasal Ranger D/T and $(D/T)_G$ (0.78 to 0.80). Correlations were higher for $(D/T)_G$ than for D/T, meaning that using the geometric mean of the unit D/T for the device provided better correlations to the other methods than using the unit D/T directly. This difference was less pronounced for the Nasal Ranger, suggesting that using geometric scale settings did not improve correlations between the Nasal Ranger data and the data from the other methods. While modest correlation (0.56 to 0.59) was found between the Nasal Ranger $(D/T)_G$ and the Mask Scentometer $(D/T)_G$, both of these methods correlated better to average intensity-predicted D/T (0.74 to 0.79 for the Nasal Ranger D/T and $(D/T)_G$, and 0.74 to 0.84 for the Mask Scentometer $(D/T)_G$). Sheffield et al. (2004a, 2004b) did not find Pearson correlation coefficients greater than 0.75 for any of the same odor methods used in our study (Mask Scentometer, Nasal Ranger, intensity, and DTFCO). However, in our study, Pearson correlation coefficients greater than 0.75 were found between the Nasal Ranger, intensity rating, and average intensity-predicted D/T.

Since correlation established the association between methods, the next step was to establish the relationships between the methods, so that by knowing one value another could be predicted. To accomplish this, linear regression was performed. Traditionally in linear regression analysis, one variable is the independent variable or predictor (x), and a relationship can be found for the response, the dependent variable (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. Therefore, in order to derive the best relationship possible

Table 2. Pearson product-moment correlation coefficients (Spearman's correlation coefficients, ϱ, in parentheses).[a]

	Intensity Rating (0-5)	Intensity-Predicted D/T	Average Intensity-Predicted D/T	Mask D/T	Mask $(D/T)_G$	DTFCO Lab D/T
Nasal Ranger D/T	0.80 (0.76)*	0.73 (0.71)*	0.77 (0.74)*	-0.22 (0.11)	--	-0.10 (0.05)
Nasal Ranger $(D/T)_G$	0.81 (0.78)*	0.77 (0.74)*	0.79 (0.76)*	--	0.59 (0.56)*	-0.10 (0.01)
Intensity rating (0 to 5)	--	0.93 (0.92)*	0.94 (0.99)*	-0.15 (0.30)	0.86 (0.84)*	0.05 (0.16)
Intensity-predicted D/T	--	--	0.98 (0.92)*	0.35 (0.35)	0.78 (0.87)*	-0.11 (0.15)
Avg. intensity-predicted D/T	--	--	--	-0.11 (0.29)	0.74 (0.84)*	-0.09 (0.15)
Mask D/T	--	--	--	--	--	-0.31 (-0.18)
Mask $(D/T)_G$	--	--	--	--	--	0.22 (0.34)

[a] Asterisks (*) indicate a significant correlation between methods ($p < 0.05$).

Table 3. Linear regression between ambient odor methods, including slopes (top values), coefficients of determination R_o^2 (middle values), and standard errors (bottom values), using session averages (*n* = 13).[a]

Independent Predictor (*x*-axis)	Dependent/Response (*y*-axis): DTFCO Lab D/T	Nasal Ranger D/T	Nasal Ranger $(D/T)_G$	Mask Scentometer D/T	Mask Scentometer $(D/T)_G$	Intensity Rating (0-5)	Intensity-Predicted D/T	Avg. Intensity-Predicted D/T
DTFCO Lab D/T	--	0.08* 0.49 0.02	0.10* 0.53 0.03	0.01* 0.28 0.005	0.02* 0.59 0.005	0.007* 0.59 0.002	0.42* 0.34 0.17	0.26* 0.43 0.09
Nasal Ranger D/T	6.3 0.49 1.8	--	--	0.10* 0.39 0.04	--	0.08* 0.92 0.007	5.72 0.80 0.8	3.29 0.87 0.4
Nasal Ranger $(D/T)_G$	5.1 0.53 1.4	--	--	--	0.19* 0.85 0.02	0.07* 0.94 0.004	4.5 0.81 0.6	2.6 0.88 0.3
Mask Scentometer D/T	28.4 0.28 13.0	3.79 0.39 1.35	--	--	--	0.37* 0.46 0.1	21.2 0.30 9.2	12.8 0.37 4.9
Mask Scentometer $(D/T)_G$	27.6 0.62 6.1	3.6 0.82 0.5	4.6 0.85 0.6	--	--	0.34* 0.94 0.02	22.8 0.83 3.0	12.8 0.86 1.5
Intensity rating (0 to 5)	76.6 0.56 18.3	10.7 0.92 0.92	13.7 0.94 1.0	1.26 0.46 0.39	2.8 0.94 0.2	--	66.6* 0.88 7.3	38.2* 0.94 2.8
Intensity-predicted D/T	0.82 0.34 0.3	0.14* 0.80 0.02	0.18* 0.81 0.03	0.01 0.30 0.006	0.04* 0.83 0.005	--	--	0.54 0.97 0.03
Average intensity-predicted D/T	1.65 0.43 0.5	0.26* 0.87 0.03	0.34* 0.88 0.04	0.03 0.37 0.01	0.07* 0.86 0.008	--	1.79* 0.97 0.09	--

[a] Asterisks (*) indicate stronger relationships based on the lowest standard error and the better model based on regression of all methods as independent and dependent variables (regressions without asterisks are shown only for information purposes). To scale Nasal Ranger $(D/T)_G$ to Mask Scentometer $(D/T)_G$, multiply by 0.19 (i.e., 1 NR = 0.19 MS). To scale a method in the gray boxes, use the inverse slope. For example, to relate Nasal Ranger $(D/T)_G$ to average intensity-predicted D/T, the stronger relationship is 0.34 (rather than 2.6, because the error is lower). Therefore, multiply the D/T by 1/0.34 = 2.9 to obtain a relative predicted D/T for intensity, or 1 NR = 2.9 times the average intensity-predicted D/T.

between methods, the relationship should be based on a predictor error that is small to negligible with respect to the response variable. Thus, the standard error of the estimate was used as the criterion for which regressor should be used. The standard error of estimate is a measure of error of prediction, i.e., the lower the standard error, the higher the precision, and the more preferred the model. Each method was regressed as both an independent variable and dependent variable relative to the other methods, and the two regression models were ranked. The model with the lowest error had the better slope or scaling factor produced from the regression. The slope values labeled with an asterisk (*) in table 3 produced the lowest error and represent the more precise relationships. The slopes and the goodness of fit of the relationships (coefficients of determinations, R_o^2, through the origin) for the session averages from the linear regression analysis are shown in table 3. Note that the R_o^2 values are the same for each of the linear models. From table 3, we can relate one method to another and assess the scale of measurements from the different methods. For illustration, the slope between the Mask Scentometer $(D/T)_G$ and the Nasal Ranger $(D/T)_G$ is about one-fifth (0.19), so the Nasal Ranger $(D/T)_G$ readings were about 5 times higher than the Mask Scentometer $(D/T)_G$ readings.

The slope for regression of two perfectly comparable methods, i.e., methods that both produce the same result, would be 1.0, and the methods would have a coefficient of determination (R_o^2) near 1.0. The coefficient of deter–mination is the proportion of the variability that is accounted by the linear model and describes the goodness of fit of the linear estimated slope. The relationship between intensity-predicted D/T and average intensity-predicted D/T is closest to a 1:1 slope at 1.79 (table 3), and the relationship is very strong ($R_o^2 = 0.97$). This good-fitting relationship is at least somewhat intuitive since both D/T values are predicted from the same set of intensity data. Other methods that showed reasonably close and strong relationships, based on this simple regression analysis, were DTFCO and intensity-predicted D/T, Mask Scentometer D/T (and $(D/T)_G$) and intensity ratings, and Nasal Ranger $(D/T)_G$ and average intensity-predicted D/T. The strongest R_o^2 values, beside the R_o^2 values between predicted D/T as just described, all involved intensity ratings, as follows: intensity vs. Mask Scentometer $(D/T)_G$ (R_o^2 = 0.94), intensity vs. average intensity-predicted D/T (R_o^2 = 0.94), intensity vs. Nasal Ranger $(D/T)_G$ (R_o^2 = 0.94), and intensity vs. Nasal Ranger D/T (R_o^2 = 0.92). The R_o^2 value between the Nasal Ranger and Mask Scentometer $(D/T)_G$ was good (0.85), as were the R_o^2 values between average intensity-predicted D/T and Nasal Ranger $(D/T)_G$ (0.88) and Mask Scentometer $(D/T)_G$ (0.85). In general, these methods have good fitting relationships between them.

Using geometric average D/T for the Mask Scentometer and Nasal Ranger improved the R_o^2 values from other methods in all instances. The slopes also came closer to 1:1 when $(D/T)_G$ was used. For example, the R_o^2 value improved from 0.39 to 0.85 between the Mask Scentometer and Nasal Ranger, and the slope increased from 0.10 to 0.19. These results are compelling for the use of $(D/T)_G$ for two reasons: first, there was a dramatic increase in accountability of the variation; and second, a high R_o^2 is essential, whereas a slope near 1 is only desirable.

In general, the relationships of laboratory DTFCO had low coefficients of determination (R_o^2 = 0.34 to 0.62). The slopes between intensity-predicted D/T (0.42) and average intensity-predicted D/T (0.26) were nearer to 1, but had low R_o^2 values (not a strong relationship). Additionally, the slopes of the Nasal Ranger, Mask Scentometer, and intensity-based predictions versus laboratory-based olfactometry (DTFCO Lab D/T) were very far from 1, requiring large scaling factors to relate DTFCO to these methods (top row of table 3), a very undesirable result.

Coefficients of determination (R_o^2) for predicted D/T were degraded slightly relative to using the intensity ratings directly, meaning that using intensity ratings to predict D/T weakened the goodness of fit. The R_o^2 values between predicted D/T and observed intensity ratings were not as good as expected, at R_o^2 = 0.88 and 0.94 for intensity-predicted D/T and average intensity-predicted D/T, respectively. In fact, the R_o^2 (0.94) for intensity ratings and the Nasal Ranger $(D/T)_G$ and Mask Scentometer $(D/T)_G$ were just as good. Perhaps something is lost in the prediction, or it is not robust. There are two schools of thought concerning the best application of the D/T prediction equation for intensity. Conceptually, it seems logical that when an assessor rates intensity, the rating corresponds directly to a predicted D/T for that assessment. Averaging the predicted D/T should then normalize the predicted D/T. The alternative is to average the series of intensity ratings for the given period of time, which has the effect of normalizing the assessment data, and then transform the intensity value to a predicted D/T. Therefore, the question became: should one normalize the raw data or the predictions? Average intensity-predicted D/T values were better correlated to the other methods (except for DTFCO Lab D/T) and had slopes closer to 1 than did intensity-predicted D/T. The prediction equation is an exponential function, so we would not expect a perfect fit to a linear model. This is the most likely reason that the exponential effect is less pronounced when the average intensity-predicted D/T is used. Again, the averaging of the intensity ratings is normalized first, and then transformed, rather than trying to fit the average of all the individual transformed assessments to a linear model. It appears from this work that using predicted D/T based on averaged intensity ratings is preferable, in terms of being better correlated to other odor assessment methods, than is averaging D/T values that were predicted from individual intensity ratings.

Table 4. Means for all measures of D/T for 13 sessions (*n* = 13 for all methods).

Method	Mean D/T[a]	SD	Session Mean Max.	Session Mean Min.
DTFCO lab D/T	134.36 a	95.6	331.0	27.7
Intensity-predicted D/T	89.00 b	78.9	290.4	7.8
Avg. intensity-predicted D/T	53.45 bc	37.6	148.8	16.1
Nasal Ranger D/T	16.20 c	8.8	31.4	4.3
Nasal Ranger $(D/T)_G$	21.10 c	9.9	35.3	6.1
Mask Scentometer D/T	2.37 c	2.0	7.1	0.5
Mask Scentometer $(D/T)_G$	4.14 c	2.2	7.4	0.5

[a] Means followed by the same letter are not significantly different at α = 0.05.

The least significant difference (LSD) multiple comparison results (table 4) showed no significant difference between the intensity-based methods and no differences between the average intensity-predicted D/T, Nasal Ranger, and Mask Scentometer data with either D/T or $(D/T)_G$. However, laboratory assessment (DTFCO) was significantly different from the other methods.

While no statistically significant differences in the session means existed between the Nasal Ranger, Mask Scentometer, and intensity-based methods, the methods did not produce the same results. The slope difference between the Mask Scentometer and Nasal Ranger may be caused by the fact that their "stops" along the D/T scale are not at the same places, the range of the Mask Scentometer is limited (0.35 to 18 D/T), and the number of assessments between methods was not the same. That is, the lower D/T for the Mask Scentometer may be a result of 20 assessments compared to two assessments from the Nasal Ranger and is likely a better representation of the room odor concentration. The researchers noted that the odor in the room decreased over the 10 min period, as the manure source equilibrated over time and less odor was generated from the source, which could explain differences between the Mask Scentometer and intensity methods compared to the other methods since these methods assessed odor during the entire session. The data were explored to substantiate this reduced odor intensity noted by the researchers, but the differences between the start and ending intensity ratings were not found to be significant. It can only be speculated that the differences between the Mask Scentometer and Nasal Ranger are due to the number of samples taken during the assessment, the different "stops" in the devices, or some other difference in the materials or physics of the devices.

If we use the Nasal Ranger $(D/T)_G$ for reference, eight of the 13 session means were higher (19.4, 22, 22.6, 24.6, 28, 32.8, 35, and 35.3) than the maximum D/T setting (18 D/T) of the Mask Scentometer (table 5). When data from only sessions 4, 6, 7, 11, and 13 (for which the Nasal Ranger

$(D/T)_G$ < 19 D/T were analyzed), the R_o^2 for Mask Scentometer $(D/T)_G$ and Nasal Ranger $(D/T)_G$ increased from 0.85 to 0.94, and the slope increased from 0.19 to 0.30 for $(D/T)_G$ and from 0.10 to 0.25 for D/T, supporting the hypothesis that the range of the Mask Scentometer is a factor in these results. This assumes that $(D/T)_G$ is equivalent between the Nasal Ranger and Mask Scentometer. Additionally, it seems logical that the Mask Scentometer would "average" out a few high D/T values, whereas just one high or low D/T from the Nasal Ranger could skew the results (only two assessments per session were taken). In addition, fewer assessors were available for Mask Scentometer readings than for intensity ratings and the Nasal Ranger; with more replication, the results could have improved. Therefore, the range of the Mask Scentometer is thought to have been a limitation. Nonetheless, from the regression analysis, a scaling factor appears to be necessary to compare Mask Scentometer results with Nasal Ranger results, and vise versa. One particularly important issue with the olfactometry data could be related to the fact that unbaked bags were used for the DTFCO method. Parker et al. (2003) found that unbaked bags had a background detectable odor of between 20 and 60 D/T, while baked bags had a range of between 12 and 16 D/T. Using unbaked bags could have contributed to the higher D/T values for DTFCO than if baked bags were used and could partially explain why the DTFCO D/T values were so much higher than the other methods. However, baking of bags is not a requirement under ASTM Standard E679-04 (ASTM, 1999b) or the European olfactometry standard (CEN, 2003).

A potential application of this work is to use it to convert the results obtained from one ambient odor method to another. This could be helpful for researchers conducting odor studies to understand how results obtained using one ambient odor method may compare to another method. The results obtained in this study are summarized as a more user-friendly conversion tool in table 6. The $(D/T)_G$ results are shown because they were found to have stronger relationships than the D/T averaging method.

The current study found the slope to be 0.08 for a Nasal Ranger and 0.01 for a Mask Scentometer, or 0.1 and 0.02 respectively, if geometric means are used. Newby and McGinley (2004) found that 7 D/T with a Nasal Ranger equated to 106 D/T using DTFCO (slope of 0.07). This agrees more with Newby and McGinley (2004) than with Brandt et al. (2008) and Bokowa (2008), who both found a factor 2 to 3 times lower with the Nasal Ranger than DTFCO (an equivalent slope of 0.3 to 0.5). An example of how to relate methods between each other using the slopes found in this study is shown in table 7. For 106 D/T using DTFCO, our slopes equate to 8 D/T and 11 $(D/T)_G$ for the Nasal Ranger, and to 1 D/T and 2 $(D/T)_G$ for the Mask Scentometer. Additionally, a Nasal Ranger $(D/T)_G$ of 7 is equivalent to a Mask Scentometer $(D/T)_G$ of 1.3, DTFCO D/T of 70, an intensity rating of 0.5 and an average intensity-predicted D/T of 18.

In this study, an intensity of 2 equates to a Mask Scentometer $(D/T)_G$ of 6, a Nasal Ranger $(D/T)_G$ of 20, an intensity-predicted D/T of 133, and a DTFCO D/T of 286 (table 7). Newby and McGinley (2004) and Huey et al. (1960) suggested that a D/T of 7 (the regulatory limit in Missouri at the time) is the threshold at which annoyance occurs. Clearly, we do not have a sound science concerning the threshold D/T that defines annoyance, but it is clear that there are distinct

Table 5. Session means for ambient odor methods.

Session	DTFCO	Intensity	Intensity D/T	Mask Scentometer D/T	Mask Scentometer $(D/T)_G$	Nasal Ranger D/T	Nasal Ranger $(D/T)_G$
1	32.0	0.6	18.6	0.8	0.5	15.7	22.6
2	76.3	2.2	136.5	3.2	3.4	29.0	32.8
3	87.7	1.7	89.3	2.4	4.6	17.3	24.6
4	99.7	1.0	33.2	5.9	2.8	8.8	12.4
5	136.3	1.9	148.5	1.3	7.4	13.7	19.4
6	63.3	0.7	21.5	7.1	1.3	4.3	6.1
7	27.7	0.8	7.8	0.7	2.4	5.9	8.2
8	59.7	2.7	290.4	2.2	6.9	27.1	35.3
9	144.0	2.0	142.1	1.8	6.3	31.4	35.0
10	197.0	1.6	66.0	1.2	4.7	15.5	22.0
11	331.0	0.8	13.6	0.5	2.2	9.0	12.6
12	208.7	1.9	98.1	2.0	6.5	22.0	28.0
13	283.3	1.7	91.4	1.8	4.8	10.9	15.4

Table 6. Suggested conversion table for relating ambient odor methods.

To Convert	Multiply by	To Obtain
DTFCO Lab D/T	0.10	Nasal Ranger $(D/T)_G$
	0.02	Mask Scentometer $(D/T)_G$
	0.01	Intensity rating
	0.42	Intensity-predicted D/T
	0.26	Avg. intensity-predicted D/T
Nasal Ranger $(D/T)_G$	10.00	DTFCO lab D/T
	0.19	Mask Scentometer $(D/T)_G$
	0.07	Intensity rating
	5.56	Intensity-predicted D/T
	2.94	Avg. intensity-predicted D/T
Mask Scentometer $(D/T)_G$	50.00	DTFCO lab D/T
	5.26	Nasal Ranger $(D/T)_G$
	0.37	Intensity rating
	25.00	Intensity-predicted D/T
	14.29	Avg. intensity-predicted D/T
Intensity rating	142.86	DTFCO lab D/T
	14.29	Nasal Ranger $(D/T)_G$
	2.70	Mask Scentometer $(D/T)_G$
	66.60	Intensity-predicted D/T
	38.20	Avg. intensity-predicted D/T
Intensity-predicted D/T	2.38	DTFCO lab D/T
	0.18	Nasal Ranger $(D/T)_G$
	0.04	Mask Scentometer $(D/T)_G$
	0.56	Avg. intensity-predicted D/T
Average intensity-predicted D/T	3.85	DTFCO lab D/T
	0.34	Nasal Ranger $(D/T)_G$
	0.07	Mask Scentometer $(D/T)_G$
	1.79	Intensity-predicted D/T

Table 7. Example method comparisons.[a]

DTFCO Lab D/T	Nasal Ranger $(D/T)_G$	Mask Scentometer $(D/T)_G$	Intensity Rating	Intensity-Predicted D/T	Average Intensity-Predicted D/T
214	15	4.5	1.5*	100	57
286	20	6	2*	133	76
50	5	1*	0.5	23	13
70	7*	1.3	0.5	32	18
106*	11	2	0.7	45	28

[a] Asterisks (*) indicate predictors used to determine other values in the same row.

differences between odor assessment methods. This work should serve as evidence that any annoyance threshold developed should be referenced to the ambient odor assessment method used to determine it.

Conclusion

In this study, dilution to threshold (D/T) results of dynamic triangular forced-choice olfactometry (DTFCO) were compared to D/T obtained using field olfactometers (i.e., the Mask Scentometer and Nasal Ranger) and results based on odor intensity ratings (using ASTM Standard E-544-99; ASTM, 1999a) under controlled conditions. The following conclusions were made:

Clearly, D/T is specific to the ambient odor assessment method with which it is measured. That is, a Mask Scentometer D/T is not the same as a D/T measured with a Nasal Ranger. When a D/T value is reported, it should be referenced to the method used to measure it. Additionally, any annoyance threshold developed should also be referenced to the ambient odor assessment method used to measure odor concentration. This has implications for regulatory limits and odor criteria in both the U.S. and other countries.

Laboratory olfactometry (DTFCO) does not correlate well with other methods when used for assessing ambient odors. The DTFCO session means were significantly different from the means for all other methods. Using intensity ratings to predict D/T (both intensity-predicted D/T and average intensity-predicted D/T) resulted in slopes nearest to 1 (0.42 for intensity-predicted D/T and 0.26 for average intensity-predicted D/T) when compared to DTFCO.

Intensity-predicted D/T values were shown to differ statistically from D/T obtained using the other odor assessment methods. Intensity ratings and average intensity-predicted D/T both correlated well to D/T obtained using the Nasal Ranger and Mask Scentometer methods. However, when an equation was used to predict D/T from odor intensity ratings, the results did not correlate as well to the other methods. Clearly, information is lost when this conversion is made.

LSD multiple comparisons showed no significant difference between the intensity-based methods ($\alpha = 0.05$) and no differences between the average intensity-predicted D/T and data obtained with the Nasal Ranger and Mask Scentometer with either D/T or $(D/T)_G$. However, laboratory assessment (DTFCO) was significantly different from the other methods. There was no statistically significant difference in the session means, even though predicted D/T based on average intensity-predicted D/T and D/T determined using the Nasal Ranger and using the Mask Scentometer were numerically noticeably different from each other. Average intensity-predicted D/T was roughly three times higher than D/T obtained using the Nasal Ranger and roughly 14 times higher than D/T obtained using the Mask Scentometer. Correspondingly, D/T obtained using the Nasal Ranger was roughly five to ten times higher than D/T obtained using the Mask Scentometer, with $(D/T)_G$ values being more similar (2 to 5 times that of the Nasal Ranger).

Results from the field olfactometry methods were more comparable to other ambient odor assessment methods when the geometric average $(D/T)_G$ was used rather than the unit D/T. In this study, using $(D/T)_G$ for the Nasal Ranger and Mask Scentometer improved the R_o^2 values (compared to D/T) between these devices and the other methods.

References

ASTM. 1999a. E544-99: Standard practices for referencing suprathreshold odor intensity. In *Annual Book of ASTM Standards.* Philadelphia, Pa.: American Society of Testing and Materials.

ASTM. 1999b. E679-04: Standard practice for determination of odor and taste thresholds by a forced-choice ascending concentration series method of limits. In *Annual Book of ASTM Standards.* Philadelphia, Pa.: American Society of Testing and Materials.

Bokowa, A. H. 2008. Odor sampling methods for point, area, fugitive, and ambient sources. Presented at the 2008 WEF and A&WMA Conference on Odors and Air Emission. Alexandria, Va.: Water Environment Federation.

Brandt, R. C., H. A. Elliot, M. A. Adviento-Borbe, E. F. Wheeler, P. J. A. Kleinman, and D. B. Beegle. 2008. Field olfactory assessment of dairy manure land application methods. In *Proc. ASABE Annual International Meeting*, ASABE Paper No. 084939. St. Joseph, Mich.: ASABE.

CEN. 2003. CEN/TC264/WG2/N222/e: Air quality–Determination of odour concentration by dynamic olfactometry; EN13725: 2003. Brussels, Belgium: European Committee for Standardization.

Henry, C. G. 2004. *Instruction Manual for Mask Scentometer.* Lincoln, Neb.: University of Nebraska-Lincoln.

Henry, C. G., G. E. Meyer, D. D. Schulte, R. R. Stowell, A. M. Parkhurst, and R. E. Sheffield. 2010. Mask scentometer for assessing ambient odors. *Trans. ASABE* 54(2): 609-615.

Henry, C. G., D. D. Schulte, R. K. Koelsch, R. R. Stowell, D. P. Billesbach, and L. Koppolu. 2005. Comparing two ambient odor assessment methods for calibrating setback estimation tools for livestock facilities. Presented at the 2005 North Carolina Animal Waste Management Symposium. Raleigh, N.C.: North Carolina State University, College of Agriculture and Life Sciences.

Henry, C. G., D. D. Schulte, R. K. Koelsch, R. R. Stowell, N. Ebrahim, A. M. Parkhurst, and D. P. Billesbach. 2006. Comparing field odor assessment methods with an atmospheric dispersion model for calibrating setback estimation tools for livestock facilities. Presented at the 2006 Workshop on Agricultural Air Quality: State of the Science. Potomac, Maryland. Sponsored by North Carolina State University, Raleigh, N.C.

Huey, N. A., L. C. Broedering, G. A. Jutze, and C. W. Gruber. 1960. Objective odor pollution control investigations. Presented at the 83rd Annual Meeting of APCP.

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). ASAE Paper No. 004044. St. Joseph, Mich.: ASAE.

Jacobson, 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 *Proc. Intl. Symp. on Gaseous and Odor Emissions from Animal Production Facilities*, 320-330. Horsens, Denmark. Sponsored by CIGR, EurAgEng, and NJF.

Koziel, J. A., J. P. Spinhirne, J. D. Lloyd, D. B. Parker, D. W. Wright, and F. W. Kuhrt. 2004. Evaluation of sample recovery of malodorous gases from air sampling bags, SPME, and sampling canisters. Presented at the 2004 ASAE/CSAE Annual Meeting. St. Joseph, Mich.: ASAE.

McGinley, M. A., and C. M. McGinley. 2003. Comparison of field olfactometers in a controlled chamber using hydrogen sulfide as the test odorant. Presented at the IWA 2nd International Conference on Odour and VOCs. London, U.K.: International Water Association.

Newby, B. D., and M. A. McGinley. 2004. Ambient odour testing of concentrated animal feeding operations using field and laboratory olfactometers. *Water Sci. and Tech.* 50(4): 109-114.

Nicolai, R. E., C. J. Clanton, and H. Guo. 2000. Modeling the relationship between threshold and intensity of swine odors. In *Proc. 2nd Intl. Conf. on Air Pollution from Agricultural Operations*, 296-304. St. Joseph, Mich.: ASAE.

Parker D. B., M. B. Rhoades, J. Koziel, and J. Spinhirne. 2003. Background odors in Tedlar bags used for CAFO odor sampling. ASAE Paper No. 034144. St. Joseph, Mich.: ASAE.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: www.R-project.org.

SAS. 2008. *Statistical Analysis Software.* Cary, N.C.: SAS Institute, Inc.

Sheffield, R. E., and P. Ndegwa. 2008. Sampling agricultural odors. PNW 595. Moscow, Ida.: University of Idaho.

Sheffield, R. E., M. Thompson, B. Dye, and D. Parker. 2004a. Evaluation of field-based odor assessment methods. Presented at the 2004 WEF and A&WMA Conference on Odors and Air Quality. Alexandria, Va.: Water Environment Federation.

Sheffield, R., M. Thompson, B. Dye, and D. Parker. 2004b. Sniff and tell: Study evaluates field-based odor assessment devices and methods for several odor sources. *Water Environ. Laboratory Solutions* 11(5): 8-11. Alexandria, Va.: Water Environment Federation.

Sheffield, R. E., M. Chahine, B. Dye, and M. Thompson. 2007. Odor concentrations downwind of Idaho dairies and heifer facilities. *Trans. ASABE* 50(1): 255-263.

Trabue, S. L., B. J. Kerr, C. J. Ziemer, and B. L. Bearson. 2008. Animal emissions analyzed by both chemical and odor panel methods. In *Proc. A&WMA Annual Meeting*. Pittsburgh, Pa.: Air and Waste Management Association.

Zhu, J., L. D. Jacobson, D. R. Schmidt, and R. E. Nicolai. 2000. Evaluation of Inpuff-2 model for predicting downwind odors from animal production facilities. *Applied Eng. in Agric.* 16(2): 159-164.