

PROTOCOLS FOR RELIABLE FIELD OLFACTOMETRY ODOR EVALUATIONS

R. C. Brandt, M. A. A. Adviento-Borbe, H. A. Elliott, E. F. Wheeler

ABSTRACT. Specific gasses (odorants) are often poorly correlated with odors, which require human perception. Thus, olfactometry is used to quantify odors, which commonly contain a complex mixture of offensive compounds. Laboratory-based dynamic olfactometry is expensive and time-consuming, and it is accompanied with sample container/ preservation issues. Field olfactometry provides real-time measurements at lower detection levels, but is influenced by environmental factors. This study explores the use of field olfactometry for quantifying dilutions-to-threshold (D/T) of environmental malodors. Nasal Ranger[®] Field Olfactometer (NRO) instruments were used to collect 3096 individual D/T observations at livestock facilities in central Pennsylvania. Twelve to 16 observations were collected at each sampling station using multiple assessors, capturing four concurrent readings each. The multiple-assessor repeat observation (MARO) technique revealed that the reproducibility of D/T observations (across assessors) was more precise than replicate observations by individual assessors (repeatability). Observations were significantly ($P < 0.0001$) influenced by odor source distance, wind direction, barometric pressure, and wind velocity. Power analysis showed that the 16-sample MARO using NRO method achieved 95% odor panel confidence with a power value of 0.90 at lower-D/T (2,4) and upper-D/T (30, 60) levels. Mid-range D/T settings of 7 and 15 exhibited the greatest odor panelist variability. This study shows that MARO field olfactometry can reliably estimate odor D/T differences, even with weather variations. It is noteworthy, however, that the greatest numbers of observations ($n = 85-91$) are needed at D/T levels of 7 to 15 (to achieve 95% confidence), precisely the range used to define nuisance odor conditions in some states.

Keywords. Field olfactometry, Livestock, odors, Multi-assessor repeat observation, Nasal Ranger[®] Field Olfactometer.

Development pressures are steadily consuming open space that once served as a buffer between agriculture and competing land uses (Greenleaf, 2000). At the same time, the number of animals kept on many farms is increasing. Odor-related nuisance complaints are on the rise, and producers are under increasing pressure to manage off-site odor impacts. This is a difficult challenge because of the lack of practical tools for quantifying odors. Nearly 300 chemical compounds have been identified in swine manure (Yin-Cheung et al., 2008). While identification of individual odorants is valuable for understanding mechanisms involved, basing decisions on odorants alone can lead to underestimation of the nuisance potential of the complex mix of malodorous gases typical of agricultural operations (Miner, 1995). Human olfactory measurement (olfactometry) offers the only reliable means for sensory quantification of odors (Miner, 1995; Zhang

et al., 2002) and therefore remains the ultimate determinant for nuisance odor episodes.

Detection threshold olfactometry strategies fall into two categories: (1) laboratory-based analysis of containerized samples; and (2) direct *in-situ* ambient air measurement methods. Both approaches involve controlled mixing of odorous air with non-odorous air to achieve known dilutions, which are presented to human subjects for evaluation. The process starts with exposure of odor panelists to a highly diluted air sample, where odor-containing air cannot be distinguished from odorless carbon-filtered air. Subjects are methodically presented with progressively lower dilution levels (greater odorous air content) in measured steps. The dilutions-to-threshold (D/T) point corresponds to an odor concentration at which the observer detects air is “no longer the same as it was before” (McGinley et al., 2000). At this dilution the observer may not be able to identify the odor, but can detect that something is different about the air relative to the previous instrument setting. The D/T is a unitless ratio calculated as:

$$D/T = \frac{\text{Volume of carbon filtered air}}{\text{Volume of odorous air}} \quad (1)$$

Detection of an odor at high dilution (high D/T) indicates the presence of a strong odor. Conversely, detection at low dilution (low D/T) indicates a relatively weak odor. Odors do not necessarily correlate well with odorant concentrations (McGinley et al., 2000), hence analytical quantification of one or more specific air-borne compounds is not a reliable indicator of nuisance odor potential.

Submitted for review in August 2010 as manuscript number SE 8738; approved for publication by the Structures & Environment Division of ASABE in February 2011. Presented at the 2009 ASABE Annual Meeting as Paper No. 09-6134.

The authors are **Robin C. Brandt**, ASABE Member Engineer, Professional Engineer, Lecturer, **M. Arlene A. Adviento-Borbe**, Project Scientist, Dept of Plant Sciences, University of California-Davis, One Shields Ave, Davis, CA 95616, **Herschel A. Elliott**, Professor, Professional Engineer, and **Eileen F. Wheeler**, ASABE Member Engineer, Engineer, Professor, Department of Agricultural and Biological Engineering (ABE), The Pennsylvania State University, University Park, Pennsylvania. **Corresponding author:** Robin C. Brandt, Department of Agricultural and Biological Engineering (ABE), The Pennsylvania State University, 101 Ag Engineering Bldg., University Park, PA 16801; phone: 814-865-2809; e-mail: rcb100@psu.edu.

Because olfactometry involves the use of a human detector, despite efforts to be objective, observations can be influenced by factors such as anxiety, distraction, disinterest, fatigue, health condition, personal comfort, or even visual cues. When outside the laboratory, the ability to manage factors that may influence odor assessors disparately is limited. Hence, laboratory-based measurement is considered to be the best available technology (Zhang, et al., 2002). While “most olfactometers can provide internally consistent results,” dynamic triangular forced-choice olfactometry (DTFCO) produces “better accuracy, reproducibility, and statistical reliability” than other methods (U.S. EPA, 1996). Today, international standard EN13725:2003, “Determination of odour concentration by dynamic olfactometry” (CEN, 2003), is widely embraced as the gold standard for laboratory-based DT (detection threshold) quantification using the DTFCO method. Unfortunately, DTFCO is expensive and time-consuming, and real-time measurement is not practical. Potential odor sample adulteration/bias introduced by the collection vessel (typically a pre-conditioned Tedlar® bag) continues to be an issue with off-site odor evaluation (Parker et al., 2003; Koziel et al., 2005; Qu and Feddes, 2006; Trabue et al., 2006). Use of DTFCO for routine odor management decision-making is rarely practical due to the need for specialized air sampling equipment/ bags, logistics associated with sample transport/ holding time, scarcity of qualified DTFCO laboratories and expense.

In the late 1950s, the U.S. Public Health Service sponsored research leading to the development of a relatively inexpensive, hand-held device for human sensory quantification of odor D/T. The first commercially available field olfactometer was manufactured by the Barneby-Sutcliff Corporation and marketed under the name, Scentometer®. The Scentometer® produces known odor dilutions by mixing ambient (odorous) air with carbon-filtered (odor free) air. The mixture is sniffed and evaluated by the odor assessor in the field. Direct on-farm D/T observations offer several advantages over laboratory DTFCO measurements, including: lower detection levels [laboratory olfactometry sensitivity is typically estimated at 20 DT to 50 DT (Gostelow et al., 2003; McGinley and McGinley, 2006; Henry et al., 2010)]; real-time readings; elimination of the need for sample collection; lower cost (McGinley and McGinley, 2003); and . convenience (Miner, 1995).

Field olfactometry proponents recommend its use as a proactive tool for monitoring routine operations; comparing operating practices; documenting odor release episodes; determining facility baseline status; and prioritizing odor sources for control measures. In some states D/T standards have been established and field olfactometry is used as a regulatory tool to verify complaints and determine compliance (Redwine and Lacy, 2000; Mahin, 2001; McGinley and McGinley, 2003; SRF Consulting Group Inc., 2004; Maine DEP, 2009).

The need for an inexpensive practical odor quantification method and recent advancements in equipment and techniques has spurred renewed interest in field olfactometry. The Nasal Ranger® Field Olfactometer (NRO, fig. 1; St. Croix Sensory, Inc., Lake Elmo, Minn.), employs several equipment and procedural modifications that facilitate collection of field D/T observations. Routine assessor odor acuity screening; instantaneous sniff flow rate

monitoring (Sheffield et al., 2004); capability for rapid replacement of spent carbon filter cartridges in the field; improved user comfort (Newby and McGinley, 2003), and; regular equipment calibration, cleaning, and inspection provide improved confidence in field D/T measurements. The ability to operate NRO units in a manner that conceals the D/T setting from the assessor also helps to avoid bias.

The growing need for cost-effective odor quantification suggests that wider use of field olfactometry in agriculture can be anticipated. Adoption of field olfactometry for nuisance odor assessments in some states provides evidence of this trend. Currently, eight states (CO, CT, IL, KY, MO, NV, ND, WY) use field olfactometry limits of 7 to 15 D/T for defining selected nuisance odor conditions (Maine DEP, 2009). In Canada, field olfactometry technology has stimulated renewed interest in its use for assessing odor complaints (Zhang et al., 2002).

Statistical evaluation of field olfactometry precision under actual field conditions in the literature is limited. Sheffield et al. (2004) investigated paired field odor assessor response variability (various odor sources) using portable olfactometry D/T devices, including box and mask Scentometers, and NROs. Laboratory-based DTFCO results for Tedlar® bag samples secured during field olfactometry observations were also evaluated for comparison. Sheffield et al. (2004) found that odor assessor response variability was lowest (greatest precision) with the NRO and DTFCO methods. This study did not assess potential environmental/ logistical factors that may have influenced odor assessor responses and observation repeatability/reproducibility were not reported. Henry et al. (2010) compared odor thresholds observed in a controlled environment chamber using Mask Scentometers, NRO units, DTFCO and Odor Intensity Reference Scale (ASTM E544-99, 2004) estimated D/T. These investigators focused on differences among observations from the various threshold techniques, working with mean assessor responses. Assessor observation repeatability/ reproducibility levels for the alternative methods were not reported. A study conducted by McGinley and McGinley (2003), using known hydrogen sulfide concentrations, found no significant difference among observations by multiple NRO assessors ($P=0.309$) in a controlled environmental chamber.

This study aimed to better understand the strengths and limitations of NRO for agricultural odor assessment, and identification of a standardized protocol for collection of high quality field olfactometry data. Thus, specific objectives were: (1) to identify and test a practical protocol for collection of *multiple assessor repeat observation* (MARO) field olfactometry sampling; (2) to quantify MARO sampling repeatability and reproducibility, (3) to identify significant environmental variables affecting MARO readings, and; (4) to determine the number of MARO samples required to achieve a 95% ($\alpha=0.05$) statistical confidence level.

METHODS AND MATERIALS

For this study, NRO units were employed to collect observations on several Pennsylvania livestock farms. Considerable logistical coordination was required because volunteer odor assessors were used and farm facilities/



Figure 1. (a) The Nasal Ranger® field olfactometer (NRO) in use. (b) D/T dilution dial located at the air intake end of the unit, which is unseen by the odor assessor during use (100% carbon-filtered air blank positions are located with arrows, between D/T labels). (c) Multiple assessor odor panel facing upwind toward odor source. (d) Multiple assessor odor panel, facing downwind from odor source. Tripod weather station set-up is in foreground.

manure management activities were under the control of others.

SELECTION OF LIVESTOCK FACILITIES

Livestock operations were located within 30-min travel distance from the Pennsylvania State University (PSU), University Park, Pennsylvania. Dairy, poultry, and swine operations were included to assure that the majority of Pennsylvania animal agriculture was represented. PSU managed livestock facilities within 10 miles of campus were selected for most field data collection. One privately owned dairy operation was included to record observations of recently field-applied dairy manure. Emissions from three livestock housing facilities (one dairy, one poultry, and one swine), two land application sites (one dairy and one swine), and one dry-stack dairy manure storage pad were evaluated over a 12-month period. Twenty-two of 29 field data collection events focused on odors originating from animal housing.

ODOR ASSESSOR RECRUITMENT, SCREENING, AND SELECTION

Seventeen volunteer odor assessor candidates were evaluated for odor sensitivity using the n-butanol “Sniffin’ Sticks” (Pen-Test) method prescribed by the NRO

manufacturer (St. Croix Sensory, 2006). It was originally intended that assessors with similar (or identical) Pen-Test scores would be used for all odor observations, and that anosmic or hypersensitive individuals would be excluded. Ultimately, due to volunteer odor assessor availability constraints, a total of 14 assessors participated; with Pen-Test scores typical ranging from 6 (slightly below normal; normal = 8) to 14 (hypersensitive). Inclusion of individuals with widely varying Pen-Test scores provided an opportunity to include this variable in statistical analyses to determine if n-butanol odor sensitivity was a significant variable affecting precision of the MARO field olfactometry technique.

Pen-testing of approved odor assessors was routinely repeated during on-farm evaluation periods. In all, ~100 Pen-tests were performed during the study (across all assessors). Testing of participating assessors was typically performed within one week of field events, with the goal of minimizing elapsed time between testing and odor panel data collection.

ON-FARM DATA COLLECTION

To prevent odor desensitization in the field, ½-face respirator masks were used [North 7700 series silicone half mask respirators with Multi-purpose (755C) North gas and vapor cartridges; Northern Safety Co., Inc., Utica, N.Y.].

Each odor assessor was required to wear a respirator mask while on-farm, except when odor characterization of whole air was being performed. These inexpensive masks proved very effective in removing ambient odors for all situations encountered in this study.

Weather observations/ measurements were recorded at each livestock facility during odor observations. Visual observations of cloud cover and precipitation were recorded, and a portable weather station was used to measure temperature, relative humidity, and barometric pressure [Kestral[®] 4000 Pocket Weather Tracker; Nielsen-Kellerman, Boothwyn, Pa.]. The portable weather station was also used to determine wind velocity for each individual D/T observation, at all stations. To accomplish this, the weather station unit was mounted with a digital clock ~1.5 m above the ground on a portable tripod. During observations, the tripod unit was placed immediately upwind (~1.5 m), between the odor source and the panel. In this way each assessor could observe the time and wind speed when they first detected odor.

The protocol used for NRO observations was based on the manufacturer's recommendations for routine use, but several additional measures were incorporated to facilitate replicate observations and reduce operator bias. For example, odor assessors were not permitted to see the D/T dial setting on their NRO unit, or those of other assessors, at any time during data collection. Each unit was set to a blank position (unknown to the NRO user) by the test administrator (TA) by rapidly rotating the dilution dial clockwise and counterclockwise, eventually stopping on a desired blank position immediately before each observation set was initiated. When the TA judged that odors at a particular field station might reach or exceed 60 D/T, the dilution dial was set at the blank position between D/T=60 and D/T=2. In all cases, odor assessors began their field olfactometer observations at a blank position and rotated the dilution dial in the direction producing decreasing D/T order. Observations were collected in sets of four, with assessors taking four individual replicate readings. Thus, 12 to 16 NRO observations were collected at each location. All assessors took each observation replicate at the same time, so that D/T detection times typically differed by <1 min. A full set of observations at each station typically took ~8 min.

Stations selected for NRO odor panel observations were chosen by the TA at each livestock facility to investigate the full range of NRO unit dilution dial (D/T) detection settings. Several stations were re-visited on more than one farm visit to collect readings under differing weather conditions. All on-farm observations were manually recorded by the TA.

PROTOCOL FOR MULTIPLE NRO ASSESSORS AND REPLICATE OBSERVATIONS

Several alternative approaches for collection of MARO field olfactometer data were explored during preliminary trials. Two preliminary trial runs employed the procedure ultimately selected for use in the study, and were therefore included in final statistical analyses. The following step-by-step procedure describes the methodology that was ultimately employed for collection of data evaluated in this study (Objective #1):

Before Going to the Field for a MARO Olfactometry Session: (1) All NRO odor assessor candidates are screened and trained, and observe EN13725:2003 (CEN, 2003)

practices regarding avoidance of perfumes, etc. during odor panel sessions; (2) The TA prepares all field data collection forms for planned odor assessment activities; (3) The TA instructs and observes NRO assessors during inspection of their respective NRO units/masks and ½-face respirators. [Inspection includes: examination for relic odors; proper functioning of NRO valves; replacement of mask comfort seals; nose piece fit; and NRO battery status.]; (4) The TA inspects the pocket weather station, digital clock, and tripod assembly for proper operation; (5) The TA prepares a farm-pack with extra batteries, NRO masks, comfort seals, ½-face respirator(s), carbon filter replacement cartridges, and; hardcopy data collection forms;

In the Field before Initiating MARO Session: (1) The TA instructs all participating NRO assessors to don their respirators to eliminate risk of odor desensitization upon arrival at the farm; (2) The TA guides NRO assessors to an observation point, evaluates wind direction, and assembles the tripod, weather station, and digital clock so that NRO users can view the tripod set-up while facing upwind (air movement into assessors faces); (3) The TA organizes NRO panelists shoulder-to-shoulder facing the tripod assembly so that all can easily see wind speed and digital time readout displays (Note: As shown in figure 1, in this arrangement, NRO units are separated by ~0.5 m during operations); (4) The TA records on-site ambient weather conditions, including cloud cover, fog, and precipitation; (5) The TA asks assessors if they can sense any odors while still wearing their respirator masks. If all assessors report no detection of odors, then the test can begin. (Note: A slight carbon-filter odor is sometimes unavoidable for some assessors.);

In Field MARO Data Collection: (1) The TA sets each NRO unit to a blank (100% carbon filtered air) setting. The blank setting is selected to be above the expected ambient air D/T level. Sometimes this means setting the dial at the blank above 60 D/T (between D/T=60 and D/T=2), which is the normal operation start position. Assessors are informed that NRO units are set at different blank positions to minimize peer pressure to report detection at the same time as other panelists; (2) The TA instructs assessors to remove their respirators and don their NRO units without breathing ambient air; (3) The TA instructs assessors to breathe through their NRO units, set at neutral (BLANK with no odor) until they are each confident that a proper mask/face seal is achieved and they are ready to begin (30 to 90 s); (4) Each assessor operates their own NRO unit at their own pace, methodically and carefully adjusting the NRO unit D/T dial settings. Assessors are instructed to make a judgment concerning odor detection after three inhalations in the target respiration zone (16 to 20 L/min) at each D/T setting in decreasing series (i.e., 60, 30, 15, ... 2) until detection (or no detection) is concluded. Dilution dial blank positions between D/T settings are used by assessors to clear their nose and pace their observations; (5) When an assessor determines a detection, he or she immediately memorizes the time and wind speed from the tripod weather station. The assessor then removes the NRO from his/her face and re-dons their respirator; (6) Assessors who complete their observation wait until the whole panel records a detection or completes the entire dilution-dial series. Odor assessors are not permitted to see the dilution-dial setting on any of the NRO units during or between observations; (7) When all odor assessors have completed their D/T observation, the TA records the

dilution-dial D/T setting on the NRO unit (when detection was observed) and the time and wind speed coinciding with detection, as stated by each assessor; (8) After all information is recorded and NRO units are re-set to new BLANK dilution-dial settings by the TA, the process is repeated (Steps 1 through 7 above). Using this procedure, 16 individual observations (4-assessors × 4-readings) can typically be obtained in ~8 min.

DATA PROCESSING AND STATISTICAL EVALUATION

All analyses in this study involved log₁₀ transformation of D/T data to convert measurements to a linear ratio-scale (Log D/T). As illustrated in figure 2, replicate observations taken by a single assessor (repeatability) and observations taken by different assessors (reproducibility) during the same session, together, reflect the precision of NRO instrument (plus operator) MARO data at a particular station. This relationship is described in equation 2:

$$\begin{aligned} \sigma_{\text{Measurement Error}}^2 &= \sigma_{\text{PO Instrument+Operator}}^2 \\ &= \sigma_{\text{Repeatability}}^2 + \sigma_{\text{Reproducibility}}^2 \end{aligned} \quad (2)$$

where σ^2 is the variance for the parameter identified in the subscript.

Expression of variance as a proportional measurement, rather than actual variance values, provides a means for interpretation routinely employed in ANOVA (regression) modeling. The term Repeatability is defined in equation 3:

$$R = \frac{\text{Variance between observers}}{\text{Variance between observers} + \text{Variance within observers}} \quad (3)$$

where R is the Repeatability coefficient, and assumes a value between 0 and 1.0. In this context, higher values of R represent lesser variability within assessors, as a majority of variance is contained between assessors. For calculation of Reproducibility, the numerator in equation 3 is replaced with Variance within observers. The sum of the Repeatability and Reproducibility coefficients is always one.

In ANOVA modeling, the coefficient of determination (r^2) quantifies the proportion of total variation explained by the

model, which is similar to the definition of Repeatability described by equation 3. Hence, similar threshold levels have been adopted here for interpretation. Low, medium, and high Repeatability (or Reproducibility) levels have been assigned range values of 0 to 0.33, 0.33 to 0.67, and 0.67 to 1.0, respectively.

Statistical evaluations related to Repeatability analysis were performed using the PROC-MIXED program (SAS, 2003). Numerous statistical models involving a variety of nesting scenarios and covariate combinations were evaluated before selecting the final model, which produced the smallest statistical Akaike's Information Criterion (AICC) value. A low AICC indicates the mixed model that best explains the data, with the minimum number of free parameters. Data for each Farm Type (dairy, poultry, and swine) were then modeled separately to determine if repeatability for the individual species differed from the composited data. Relationships of measured environmental variables and factors affecting odor threshold estimates were calculated using Pearson correlation analysis.

Statistical power analysis was conducted to determine the number of field observations necessary to identify significant difference at a 95% confidence ($\alpha=0.05$) among different odor sources/ sites using the SAS program (SAS, 2003).

RESULTS AND DISCUSSION

In all, 32 on-farm odor panel sessions were performed to collect NRO observations at local livestock facilities. The first six were conducted at PSU swine housing and land application sites, where swine manure was surface-applied on grass hay. Preliminary statistical evaluations of these data were used to refine on-farm protocols for subsequent data collection. Ultimately, three of the first six on-farm sessions were eliminated from the final data set. Data from 29 on-farm odor panel sessions incorporating 3096 individual NRO observations were included in the study reported here.

Dilution-to-Threshold data were approximately equally divided among the three Farm-Types: dairy (1008 observations), poultry (1016 observations), and swine (1072 observations). All on-farm sessions were designed to employ four NRO assessors, but on six occasions field work proceeded with only three assessors when a volunteer unexpectedly withdrew from participation for that day and a backup assessor could not be scheduled. In all, >200 MARO data sets were collected over approximately 12 months (Fall/Spring/Summer), spanning a wide range of weather conditions.

FIELD OLFACTOMETER REPEATABILITY/ REPRODUCIBILITY

Extensive analysis of variance (ANOVA) for the full data set including all farms found Log D/T observation set Repeatability and Reproducibility coefficients of 0.27 and 0.73, respectively. Thus, with control of random effects (farm, day, and location) and fixed factors/ covariates [Farm Type, assessor odor sensitivity based on Pen-Score, Wind Speed, Wind Direction, Distance, and interactions among these factors], only 27% of the variance in Log D/T observations were explained by differences between odor panel assessors, while the remaining 73% of variance was reflected in individual assessor responses in the course of replicate observations. Odor panelist Pen-Score and Distance

Possible Discrete Field Olfactometer D/T Values						
	60	30	15	7	4	2
Location xyz		Odor Assessor ID				
		A	B	C	D	
Replicate Observation Number	1	15	7	15	15	
	2	7	15	15	15	
	3	15	15	15	7	
	4	30	30	30	30	

Repeatability (0.27)
Precision of replicate observations from individual odor assessor.

Reproducibility (0.73)
Precision of observations by different individuals comprising the odor panel.

Figure 2. Illustration of Repeatability vs. Reproducibility for a Multiple-Assessor Repeat Observation data set. [Overall, D/T levels for the full data set show less variance across assessors (reproducibility = 0.73) than replicate observations by individual assessors (repeatability = 0.27).]

from the primary odor source were highly significant ($P < 0.0001$). Farm Type ($P = 0.034$) and Wind Speed ($P = 0.037$) were significant covariates. As expected, assessor Pen-Score and Wind Speed were positively correlated with Log D/T observations, while Distance from the primary odor source was inversely related with Log D/T.

Multi-collinearity interactions among covariates complicated the analysis. Notwithstanding, results showed little variation while a number of statistical model designs were examined (not included herein). Repeatability was also assessed independently for each of the livestock species. Table 1 provides a summary of the Repeatability analysis results, which consistently reveal LOW Repeatability and HIGH Reproducibility for the composited data, and for each Farm Type.

ODOR PANEL DESENSITIZATION

A single on-farm MARO odor panel session typically lasted up to 2 h. It is reasonable to speculate that odor assessors might begin to lose focus and even become desensitized to ambient odors over time. However, no statistical evidence supporting this assertion was found. Assessor observation time in the field had no notable effect on reported D/T readings. Odor assessors who scored highest on the Pen-Test generally reported higher D/T values than individuals with low Pen-Scores. This finding is consistent with expectations, since individuals with more acute odor sensitivity would be capable of detecting odor at lower concentrations (i.e. higher D/T).

ODOR PANELIST SENSITIVITY

Odor panelist Pen-Score was found to be a highly significant variable in the full 3096 observation data set. Hence, another series of ANOVA analyses were performed with a trimmed data set, in which D/T data from

assessors with below average odor sensitivity were excluded. This analysis produced similar Repeatability results and showed Distance ($P < 0.0001$), Wind Direction ($P < 0.0001$), and Wind Speed ($P < 0.02$) were significant variables influencing Log D/T estimates. Since exclusion of assessors with poor odor sensitivity is readily accomplished by excluding low Pen-Score panelists, several additional statistical analyses were performed using a trimmed data set.

FIELD OLFACTOMETER D/T OBSERVATIONS AND ENVIRONMENTAL/SITE FACTORS

Weather conditions during on-farm MARO odor assessments are summarized in table 2. No statistical significance relating Log D/T with cloud cover, precipitation, temperature, relative humidity, or barometric pressure was found with the full 3096 observation data set. Assessors reported lower D/T observations when Wind Direction varied so that, at times, natural air movement was not transporting emissions from the primary odor source to the observation station. Detailed correlation analyses of Log D/T values from the trimmed data set were also performed for a variety of environmental/ site factors. As shown in table 3, distance, wind direction, barometric pressure, and wind speed were revealed as highly significant ($P < 0.0001$) variables influencing the estimation of Log D/T, accounting for 23.5%, 23.1%, 17.4%, 7.0% of variance, respectively. Of these parameters, the significance of barometric pressure was unexpected. In this study, barometric pressure ranged from 723 to 735 mmHg. The positive correlation coefficient indicates that odor sensitivity (Log D/T) increased with increasing atmospheric pressure. However, caution is warranted in embracing the barometric pressure/odor linkage too literally in the present study, as the relatively small pressure range in our data seems too limiting to be a primary driving force for significant D/T effects. It is possible that the

Table 1. Repeatability and Reproducibility of NRO observations^[a].

Data Set Evaluated	Calculated Repeatability	Calculated Reproducibility
Full data set ($n = 3,096$)	0.27	0.73
Dairy ($n = 1,008$)	0.32	0.68
Poultry ($n = 1,016$)	0.22	0.78
Swine ($n = 1,072$)	0.24	0.76

^[a] All assessors, regardless of Pen-Test score, are included in these data sets.

Table 2. Selected weather conditions encountered during study.

	Wind Speed ($m\ s^{-1}$)	Temperature ($^{\circ}C$)	Relative Humidity (%)	Barometric Pressure (mmHg)
Mean	1.4	21.9	41.2	729.4
Median	1.1	22.3	37.5	729.0
Mode	0.0	23.1	50.0	732.0
Min	0.0	8.2	19.5	723.1
Max	8.6	34.2	78.8	735.3

Table 3. Correlation analysis of Log D/T versus significant variables for assessors with Pen-Scores ≥ 8 (trimmed data set excluding Pen-Scores < 8).

Most significant variables ($P < 0.02$) for the estimation of Log D/T			
Variable	Description	Coefficient	P-Value
Distance	Distance between the primary odor source and observation station	-0.235	<0.0001
Wind direction	Wind not consistently originating from primary odor source throughout duration of odor panel D/T observation set	-0.231	<0.0001
Barometric pressure	Barometric pressure	0.174	<0.0001
Wind speed	Wind velocity	0.070	<0.0001
Elapsed time	Total elapsed time for odor panel to perform a full set of observations	0.068	0.0002
Pen-Score	Odor assessor sensitivity, as determined using the 'Sniffin-Stiks' n-butanol Pen-Test procedure	0.049	0.0066
Time code	Individual assessor observation replicate (1 through 4)	0.044	0.0134

influence of barometric pressure may have been unintentionally biased by odor panel timing and location. In this case, barometric pressure may actually be a surrogate indicator for one or more other combined environmental or logistical factors not evaluated in this study, such as atmospheric stability and/or vertical temperature gradient.

MULTIPLE-ASSESSOR REPEAT OBSERVATION FIELD OLFACTOMETRY DATA PRECISION

Evaluation of the full data set across consecutive replicate assessor D/T observations showed Log D/T coefficient of variation [CV = (Log standard deviation) / (Log mean)] values ranging from zero to 1.42 (table 4). Interestingly, CV levels < 0.40 were found for all individual odor panelists, regardless of Pen-Score. Table 4 shows that odor panelist CV increased with Distance from the odor source up to ~60 m, and decreased thereafter. Mean and median CVs ranged from 0.19 to 0.41 for all Distance categories. Odor panel D/T readings were likely to be more precise near the source, where greater D/T levels are observed, and then again at more remote distances, where greater dilution and low D/T levels occur. It is reasonable to speculate that better replicate observation precision would occur at the extremes of the NRO unit range, as readings of barely detectable (low D/T) or very intense (high D/T) would be more consistently identified by odor panelists, compared to middle-range D/T levels (e.g. 7 and 15 D/T).

The CV levels found in this study compare favorably with those reported elsewhere for manure odor studies employing sensory assessment. Due to the scarcity of published field olfactometry research, laboratory-based DTFCO olfactometry research provides the bulk of information for comparison. Sheffield et al. (2004) found Log D/T CVs of 0.37 to 0.47 for dairy and beef NRO observations using two-person odor assessor teams. Concurrent 15-min composited Tedlar® bag whole-air samples evaluated via laboratory-based DTFCO olfactometry consistently found mean Log Detection Threshold (Log DT) levels approximately twice the mean Log DT from NRO observations. These investigators reported greater precision with laboratory DTFCO testing, with Log DT CVs ranging from 0.17 to 0.28 (Sheffield et al., 2004). Lim et al. (2001) found laboratory DTFCO Log DT CV levels of 0.04 to 0.19 in Tedlar® bag samples of exhaust air from a swine manure pit. However, in another study of swine manure lagoon odors, Lim et al. (2003) found Log DT CVs ranged from 0.34 to 0.37, again using laboratory DTFCO odor quantification. Heber et al. (2002) reported Log DT CVs in the range of 0 to 0.35 for whole-air samples collected using a buoyant convective flux chamber, subsequently evaluated via DTFCO olfactometry. Lau et al. (2003) evaluated swine

manure odors from alternative land application techniques using a surface isolation flux chamber. In this study, DTFCO assessment of whole-air samples found Log DT CV values ranging from 0.59 to 0.88 following field application (Lau et al., 2003). Similarly, Log DT CVs of 0.88 and 0.97 were found in DTFCO results reported by Zhang et al. (2007), who used a wind tunnel to secure whole-air samples from a swine manure storage. Parker et al. (2005) studied odor emissions from beef cattle feedyards over a 12-month period, finding CV levels of 0.99 to 1.16 for pooled DTFCO Log DT data from whole-air samples. The aforementioned studies show that odor panel evaluation of containerized whole-air samples can be variable even when using laboratory-based DTFCO olfactometry. Notwithstanding, sensory assessment remains the best method for quantifying nuisance odor emissions. This present study indicates that MARO field olfactometry is capable of replicate data precision comparable with off-site DTFCO containerized sample assessment.

POWER ANALYSIS DETERMINATION OF FIELD OLFACTOMETRY SAMPLING REQUIREMENTS

Statistical power analysis was used to estimate the number of observations needed to detect a treatment difference (SAS, 2003). This is dependent on data variability (e.g. CV). High variance among observations translates into lesser ability (low power) to detect treatment differences. Low variance (high repeatability) in observations yields greater statistical power. A power value of 0.9 indicates a 90% chance the investigator will correctly infer a significant treatment effect has occurred. With a higher power value confidence improves, thus a power value of ≥0.80 is typically recommended, but 0.90 is more commonly used.

The trimmed Log D/T data set provided an opportunity for retrospective power analysis of the NRO procedure to investigate the number of observations necessary to obtain 95% confidence in results. Several CV levels up to 1.5 were evaluated, but the focus of this work settled on CVs of 0.25 and 0.50, as the data clearly shows this range is reasonably achievable, and could potentially be improved upon with improved odor panelist and odor station selection procedures (e.g. relative to odor source distance and plume vector). As shown in table 5, the number of NRO samples required to achieve a power level of >0.90 with 95% confidence, when the NRO unit setting equals 2 D/T, ranges from 3 to 4 observations (CV=0.25 and 0.50). At a setting of 4 D/T, 5 to 12 observations would be required for CVs of 0.25 and 0.50, respectively. When the NRO setting was 30 D/T, 18 observations resulted in a power value of 0.91 with a CV of 0.50. At the highest field olfactometer setting of 60 D/T, 10 observations at CV=0.50 produced a power value of 0.93.

Table 4. Field olfactometer odor assessor Log D/T coefficient of variation vs. distance between the primary odor source and odor panel observation station.

CV ^[a]	Distance between Primary Odor Source & Observation Station (m)				
	0 to 8	8 to 30	30 to 60	60 to 90	>90
Mean	0.289	0.347	0.412	0.407	0.235
Median	0.190	0.256	0.321	0.371	0.134
Mode	0.00	0.00	0.00	0.00	0.00
Minimum	0.00	0.00	0.00	0.00	0.00
Maximum	1.269	1.323	1.421	0.98	1.176

[a] CV = Coefficient of variation.

Table 5. Power analysis results showing the number of NRO observations required to achieve 95% confidence for the various D/T levels vs. data set CV and minimum power value of 0.90.

D/T	CV = 0.25 ^[b]		CV = 0.50	
	N ^[c]	Power	N	Power
2	3	0.971	4	0.907
	4	>0.999	7	>0.999
4	5	0.923	12	0.922
	9	>0.999	25	>0.999
7	25	0.906	85	0.901
	57	>0.999	203	>0.999
15	27	0.909	91	0.900
	61	>0.999	212	>0.999
30	7	0.934	18	0.910
	13	>0.999	40	>0.999
60	5	0.968	10	0.928
	8	>0.999	20	>0.999

[a] CL = Confidence level.

[b] CV = Coefficient of variation.

[c] N= Observations required for CV & power value shown.

Thus, lower range NRO levels of 2 D/T and 4 D/T, and higher range NRO levels of 30 D/T and 60 D/T resulted in logistically manageable sample numbers. The number of observations required to achieve a power level of 0.90 and 95% confidence at a CV of 0.25 for settings of 7 D/T and 15 D/T jumped to 25 and 27, respectively. Moreover, at a CV of 0.50, at least 85 observations would be needed to achieve this power level for 7 D/T, and 91 observations would be required for a 15 D/T. These results indicate that odor panel NRO readings in the range of 7 D/T to 15 D/T show the greatest odor panel variability, and thus require the greatest number of observations to obtain reliable results. This is a noteworthy finding since 7 D/T is used as a threshold for defining nuisance odor conditions in some states (Hamel et al., 2004).

The greater number of MARO observations needed to achieve 95% confidence for mid-range NRO D/T settings may be explained by the natural tendency for assessors to agree when odors are very strong (e.g. D/T levels of 30 and 60) or relatively weak (e.g. D/T levels of 2 and 4). Odor assessors agree less frequently (greater variance) at mid-range D/T levels of 7 and 15 and thus a greater number of observations are necessary.

In the present study, a practical limit of 16 observations was selected (four assessors × four replicate observations), based on personnel logistics (availability, transportation, etc.) and a desire to capture all observations within a relatively brief time (<10 min). Thus, a second power analysis was performed, holding the number of NRO observations at a constant level of 16. Table 6 summarizes the statistical confidence (percentage) that was achievable when 16 NRO observations were used with CV levels of 0.25 and 0.50. As shown, for 2 D/T, 4 D/T, 30 D/T, and 60 D/T, 16 observations produced a power score ≥0.90 with 95% confidence, for CV levels up to 0.50. When the NRO was set to 7 or 15 D/T and a CV value of 0.25 was used (with a power level of 0.9), only ~80% confidence was achievable. Furthermore, at 95% confidence and CV=0.25, the power value dropped to ~0.70 with 16 observations for 7 and 15 D/T. When the CV was set to 0.50 and 16 observations, power

scores dropped to <0.27 at 95% confidence (table 6). This evaluation shows that detection of statistically significant differences among treatments with NRO odor panel observations in the 7 to 15 D/T range are subject to the greatest uncertainty, which unfortunately corresponds with the range most commonly used in states that employ field olfactometry for regulatory decision-making for nuisance odor episodes.

CONCLUSIONS

- A practical protocol for collection of Multiple-Assessor Repeat Observation (MARO) field olfactometry data was developed. The procedure involves collection of field olfactometry D/T observations in sets of 16, using four assessors taking four consecutive readings each. MAROs were performed over a period of <10 min with panelists standing shoulder-to-shoulder facing up-wind toward the primary odor source. Field olfactometer dilution dial settings were concealed at all times to reduce assessor bias. A test administrator recorded weather and D/T data, and reset each NRO unit to a blank dilution dial position after each assessor observation. In addition to the detailed MARO protocol, a series of recommendations are offered (see appendix) to supplement manufacturer's instructions to facilitate field olfactometry reliability.
- Statistical evaluation of the full observation data set for all Farm Types (dairy, poultry and swine) revealed a Repeatability coefficient of 0.27 and Reproducibility coefficient of 0.73. Thus, most variance in odor panel D/T observations was due to dissimilar replicate readings by individual assessors over time, rather than differing readings between assessors. Similar results were produced for each livestock type when evaluated separately.
- No statistically significant relationships between NRO Log D/T and cloud cover, precipitation, temperature, relative humidity, or barometric pressure were found with the full 3096 observation data set. When odor assessors with below average Pen-Test odor sensitivity were

Table 6. Power analysis results showing the percent confidence possible with 16 NRO observations for the various D/T levels vs. data CV and resulting power value.^[a]

D/T	CV = 0.25		CV = 0.50	
	CL%	Power	CL%	Power
2	95	>0.999	95	>0.999
	95	>0.999	95	0.980
4	95	0.724	95	0.266
	90	0.832	90	0.389
	85	0.884	85	0.477
	80	0.916	80	0.545
15	95	0.693	95	0.251
	90	0.807	90	0.371
	85	0.865	85	0.457
	80	0.900	80	0.526
30	95	>0.999	95	0.870
60	95	>0.999	95	0.995

[a] CL% = Confidence level %; CV = Coefficient of variation; N = Number of observations.

excluded, distance (from odor source), wind direction, barometric pressure, and wind speed were found to be highly significant ($P < 0.0001$) variables influencing reported Log D/T values. Odor panelists reported lower D/T when wind direction varied from the vector connecting the primary odor source with odor observation station. Field olfactometer D/T values were significantly ($P = 0.0066$) influenced by assessor sensitivity. As expected, Pen-Score and wind speed were positively correlated with D/T readings, while odor source distance was inversely related with D/T. It is recommended that future studies consider incorporation of solar radiation and atmospheric stability classification to assess possible correlations with observed field olfactometry D/T levels.

- Statistical power analysis showed that the 16 sample MARO NRO method achieved 95% odor panel confidence with a power value of 0.90 at lower-D/T (2,4) and upper-D/T (30, 60) levels. Mid-range D/T settings of 7 and 15 exhibited the greatest odor panelist variability. This study shows that MARO field olfactometry can reliably detect odor D/T differences in ambient air. However, caution is warranted when mid-range D/T concentrations are observed. Nuisance odor standards that specify field olfactometry D/T levels of 7 to 15 for regulatory action may be vulnerable, especially when used for high-impact decisions.

Field olfactometry can be a valuable management tool to aid producers and agricultural advisers in decisions impacting the odor potential of production units and practices, and in evaluating odor reduction strategies. Meaningful results are contingent upon strict methodological protocols. Instruments must be frequently calibrated and used in accordance with established procedures. The suitability of odor panel members requires ongoing surveillance. Multiple odor assessors and observations should be employed when costly decisions are involved.

ACKNOWLEDGEMENTS

This work was sponsored, in part, by a USDA Special Research Grant for Improving Dairy Management Practices, and a Pennsylvania Department of Agricultural Research Grant. The authors also gratefully acknowledge the following individuals for their contributions in making this work possible: Paul H. Heinemann, Adam C. Brandt, Deborah A. Topper and Pat A. Topper, with the Department of Agricultural and Biological Engineering at the Pennsylvania State University (PSU), and; J. Zhu and Durland Shuman, with the Department of Statistics at PSU.

REFERENCES

ASTM. 2004. E544-99: Standard practices for referencing supra-threshold odor intensity. *Annual Book of ASTM Standards*. Philadelphia, Pa.: American Society of Testing and Materials.

CEN (Committee for European Normalization). 2003. EN13725: Air quality – determination of odour concentration by dynamic olfactometry. Brussels, Belgium.

Gostelow, P., P. Longhurst, S. A. Parsons, and R. M. Stuetz. 2003. Sampling for the measurement of odours. Scientific and Tech. Report No. 17. London, UK: IWA Publishing.

Greenleaf, C. 2000. Fading farmland. *American Vegetable Grower* 48(1): 61-64.

Hamel, K. C., L. Walters, C. Sulerud, and M. A. McGinley. 2004. Land application odor control case study. Presented at: *Water Environment Federation Residuals and Biosolids Management Conference*. Salt Lake City, Utah: 22-25 February 2004. Alexandria, Va.: WEF.

Heber, A. J., J. Q. Ni, and T. T. Lim. 2002. Odor flux measurements at a facultative swine lagoon stratified by surface aeration. *Applied Eng. in Agric.* 18(5): 593-602.

Henry, C. G., D. D. Schulte, S. J. Hoff, L. D. Jacobson, and A. M. Parkhurst. 2010. Comparison of ambient odor assessment techniques in a controlled environment. In *Proc.: Intl. Symposium on Air Quality and Manure Management for Agriculture*. St. Joseph, Mich.: ASABE.

Koziel, J., J. Spinhirne, J. Lloyd, D. Parker, D. Wright, and F. Kuhrt. 2005. Evaluation of sample recovery of malodorous livestock gases from air sampling bags, solid phase microextraction fibers, Tenax TA sorbent tubes, and sampling canisters. *J. Air Waste Manage. Assoc.* 55(8): 1147-1157.

Lau, A., S. Bittman, and G. Lemus. 2003. Odor measurements for manure spreading using a subsurface deposition applicator. *J. Environ. Science and Health* B38(2): 233-240.

Lim, T. T., A. J. Heber, J. Q. Ni, A. L. Sutton, and D. T. Kelly. 2001. Characteristics and emission rates of odor from commercial swine nurseries. *Trans. ASAE* 44(5): 1275-1282.

Lim, T. T., A. J. Heber, J. Q. Ni, A. L. Sutton, and P. Shao. 2003. Odor and gas release from anaerobic treatment lagoons for swine manure. *J. Environ. Qual.* 32(2): 406-416.

Mahin, T. D. 2001. Comparison of different approaches used to regulate odors around the world. *Water Sci. and Tech.* 44(9): 87-102.

Maine DEP (Department of Environmental Protection). 2009. Report on odor and gas management at solid waste facilities. Document #207-287-7718. Augusta, Maine: DEP, Division of Solid Waste Management.

McGinley, C. M., and M. A. McGinley. 2006. An odor index scale for policy and decision making using ambient and source odor concentrations. In *Proc.: WEF/AWWA Odors and Air Emissions Conference*. Alexandria, Va.: WEF.

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 Intl. Water Association, 2nd Intl. Conf. on Odour and VOCs: Measurement, Regulation, and Control Techniques*. Singapore: IWA.

McGinley, C. M., T. D. Mahin, and R. J. Pope. 2000. Elements of successful odor/odour laws. Presented at: *WEF Odor/VOC 2000 Specialty Conference*. Alexandria, Va.: WEF.

Miner, R. M. 1995. A review of the literature on the nature and control of odors from pork production facilities: An executive summary. Report prepared for the National Pork Producers Council. Corvallis, Oreg.: Bioresource Engineering Department, Oregon State University.

Newby, B. D., and M. A. McGinley. 2003. Ambient odor testing of concentrated animal feeding operations using field and laboratory olfactometers. Presented at: *The International Water Association, 2nd Intl. Conf. on Odour and VOCs: Measurement, Regulation, and Control Techniques*. Singapore: IWA.

Parker, D., M. 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.

Parker, D. B., M. B. Rhoades, G. L. Schuster, J. A. Koziel, and Z. L. Perschbacher-Buser. 2005. Odor characterization at open-lot beef cattle feedyards using triangular forced-choice olfactometry. *Trans. ASABE* 48(3): 1527-1535.

Qu, G., and J. J. R. Feddes. 2006. Estimation of measureable error caused by background odor in sampling bags. ASABE Paper No. 064141. St. Joseph, Mich.: ASABE.

- Redwine, J., and R. Lacey. 2000. A summary of state odor regulations pertaining to confined animal feeding operations. In *Proc. of the 2nd Intl. Conf. on Air Pollution from Agricultural Operations*. St. Joseph, Mich.: ASAE.
- SAS. 2003. *SAS/STAT User Guide*. Version 9.1. Cary, N.C.: SAS Institute.
- Sheffield, R., M. Thompson, B. Dye, and D. Parker. 2004. Evaluation of field-based odor assessment methods. In *Proc. Water Environment Federation and the Air and Waste Mgmt. Association Conf. on Odor and Air Quality*. Alexandria, Va.: WEF.
- SRF Consulting Group Inc. 2004. A review of national and international odor policy, odor measurement technology and public administration. Available at: www.pca.state.mn.us/publications/p-gen2-02.pdf. Accessed 8 July 2010.
- St. Croix Sensory. 2006. Odor sensitivity test. Available at: www.nasalranger.com/Operations/TP%202000%200608V2.2.pdf. Accessed 8 July 2010.
- Trabue, S. L., J. C. Anhalt, and J. A. Zahn. 2006. Bias of Tedlar[®] bags in the measurement of agricultural odorants. *J. Environ. Qual.* 35(5): 1668-1677.
- U.S. EPA. 1996. Swine CAFO odors: Guidance for environmental impact assessments. EPA Region 6. Contract No. 68-D3-0142. Dallas Tex.: Lee Wilson and Associates.
- Yin-Cheung, M. L., J. A. Koziel, L. Cai, S. J. Hoff, W. S. Jenks, and H. Xin. 2008. Simultaneous chemical and sensory characterization of volatile organic compounds and semi-volatile organic compounds emitted from swine manure using solid phase micro-extraction and multidimensional gas chromatography-mass spectrometry-olfactometry. *J. Environ. Qual.* 37(2): 521-534.
- Zhang, Q., J. Feddes, I. Edeogu, M. Nyachoti, J. House, D. Small, C. Liu, D. Mann, and G. Clark. 2002. Odour production, evaluation and control. Final report submitted to: Manitoba Livestock Manure Management Initiative, Inc. Project MLMMI 02-HERS-03.
- Zhang, Q., X. J. Zhou, N. Cicek, and M. Tenuta. 2007. Measurement of odour and greenhouse gas emissions in two swine farrowing operations. *Can. Biosys. Eng.* 49 (6): 13-20.

APPENDIX 1.

Additional recommendations to facilitate field olfactometry reliability:

- (1) Assess participant n-butanol odor sensitivity as close to the date of field work as practical (same day if possible);
- (2) Use properly fitted respirators to avoid odor fatigue before and between Nasal Ranger[®] field olfactometry (NRO) observations;
- (3) Limit repeated or extended NRO observations when air temperatures are below ~10°C (~50°F). (At low temperatures, moisture tends to condense inside respirators and NRO units. Some assessors also experience congestion from repeated inhalation of cold air);
- (4) Limit repeated or extended NRO observations when air temperatures are above ~32°C (~90°F). (Some assessors experience claustrophobia in hot weather leading to distraction that may influence reported D/T results);
- (5) Maintain NRO batteries near full charge. (When NRO batteries run low it is difficult to achieve the target inhalation flow rate. When participants experience this problem, the first corrective measure is to replace batteries);
- (6) Initiate multiple assessor repeat observation measurements at the same time. (This practice is observed to minimize the elapsed time among D/T readings across the panel);
- (7) Three sniffs of ~3 s duration each, at each dilution dial setting, are recommended until detection is observed. (This measure is suggested so that assessors can assure that mask leakage is not the source of odor detection);
- (8) Minimize odor panel distractions to the maximum extent practical. (The test administrator must be diligent to assure that all odor assessors remain focused.)