To:	Mark Dubin
From:	Jon Harcum and Steve Dressing
Date:	March 7, 2017
Subject:	CBP Technical Support: Producer Survey Evaluations

Multiple methods exist to document the extent of non-cost-shared annual and multi-year structural best management practices (BMPs) as identified by the Chesapeake Bay Program Partnership's publication entitled "Strengthening Verification of Best Management Practices Implemented in the Chesapeake Bay Watershed: A Basin Framework."[1] This technical memorandum provides an overview of a procedure that could be used to evaluate a self-certified assessment inventory (e.g., mail-in survey, online survey, etc.) that includes follow-up in-person verification using a stratified random sample of the returned producer surveys. The procedures described here could be extended to address follow-up independent verification that uses alternative sampling strategies for selecting surveys to verify.

This technical memorandum does not address selection of an appropriate survey tool (e.g., online versus mail-in), but the method described here can be used to evaluate any survey that meets the criteria described in the Summary and Discussion.

This technical memorandum assumes that independent field verification yields the truth about the presence or absence of BMPs, as well as their operation and maintenance. The specific methods for assessing the presence or absence of BMPs are not addressed by this technical memorandum.

1.0 Background

Penn State University (PSU) and the Pennsylvania Department of Environmental Protection (DEP) undertook an agricultural conservation practice inventory (survey) to capture data on visual and non-visual non-cost-share BMPs for reporting and crediting in the Bay model. The survey methodology is described in briefing materials (DEP 2016b) and a methodology report (PSU 2016). In essence, farmers were mailed conservation practices inventory forms to complete and submit to PSU. Approximately 10 percent of the responses from each county were then randomly selected for field verification by trained PSU Extension agents. Results from farmer inventories were compared against in-field independent inventories to assess the accuracy of the method. Of an estimated 33,610 farms in Pennsylvania's portion of the Chesapeake Bay watershed, PSU sent inventories to approximately 20,000 farms. A total of 6,782 surveys were returned (34%) and approximately 10 percent of the responses (710 farms) were selected for on-site verification.

PSU concluded that there was no systematic over- or under-reporting for nearly all BMPs (Royer 2016). The exceptions to this are barnyard acres under E&S/NRCS plans and riparian buffers. These practices both showed systematic over-reporting. Because their analysis showed that the over-reporting of these particular practices is statistically significant, PSU believes that an

[1] http://www.chesapeakebay.net/publications/title/strengthening_verification_of_best_management_practices_implemented_in_the

adjustment factor could be applied to adjust the cumulative dataset downward. PSU also believes that the systematic over-reporting of riparian buffer acres may be attributed to differences between the way the questions were asked in the farm survey regarding buffers and stream bank fencing, and how Extension agents were trained to record these answers during the on-farm visits.

Previously (Tetra Tech, 2016) reported on basic measures of statistical accuracy using proportion correct (PC), hit rate (HR), and false alarm rate (FAR) (see Table 1). The relatively high PC across most practices can be attributed to the large fraction of surveys where it was verified that the operation correctly reported that a practice was not in use. HR and FAR were more varied. This technical memorandum extends Tetra Tech's previous analysis to include a procedure that can be used to estimate state- and county-level acreages after adjusting for survey verification.

Table 1. Measures of survey accuracy.

Practice	Subcategory	Proportion Correct (PC)	PC Range at 90% Confidence Level	Hit Rate (HR)	False Alarm Ratio (FAR)
	Row Crops	0.85	83-87%	0.77	0.13
	Pasture Acres	0.81	78-83%	0.62	0.19
	Hay Acres	0.80	78-82%	0.67	0.24
	Privately Funded Act 38 Row Crop Acres	0.93	92-95%	0.26	0.46
	Privately Funded Act 38 Pasture Acres	0.94	92-95%	0.14	0.60
	Privately Funded Act 38 Hay Acres	0.93	92-95%	0.09	0.69
	Privately Funded NRCS 590 Row Crop Acres	0.95	94-96%	0.21	0.68
	Privately Funded NRCS 590 Pasture Acres	0.97	96-98%	0.24	0.71
	Privately Funded NRCS 590 Hay Acres	0.95	94-97%	0.23	0.75
	Manure Management Plans on Row Crop Acres	0.84	82-86%	0.61	0.39
	Manure Management Plans on Pasture Acres	0.84	82-86%	0.49	0.40
	Manure Management Plans on Hay Acres	0.85	83-87%	0.60	0.43
	Advanced Nutrient Management	0.83	81-86%	0.35	0.69
E&S Plans	Row Crop Acres	0.90	89-92%	0.30	0.46
	Pasture Acres	0.92	91-94%	0.30	0.48
	Hay Acres	0.93	91-94%	0.27	0.44
	Barnyard Acres	0.96	94-97%	0.17	0.73
NRCS Plans (privately funded)	Row Crop Acres	0.81	79-84%	0.35	0.57
	Pasture Acres	0.86	84-88%	0.28	0.58
	Hay Acres	0.85	83-87%	0.31	0.58
	Barnyard Acres	0.94	92-95%	0.16	0.78
Stream Bank Fencing	Fencing Length (Ft.)	0.88	86-90%	0.71	0.15
	Distance from Stream to Fence (Ft.)	0.87	86-89%	0.74	0.19
	Public Funded Fencing (Ft.)	0.93	92-95%	0.69	0.25
	Privately Funded Fencing (Ft.)	0.87	86-89%	0.53	0.30
	Acres of Buffer	0.87	85-89%	0.70	0.19
	Acres of Privately Funded Buffer	0.87	85-89%	0.53	0.34

Practice	Subcategory	Proportion Correct (PC)	PC Range at 90% Confidence Level	Hit Rate (HR)	False Alarm Ratio (FAR)
Riparian Buffers	Buffer Acres	0.71	68-73%	0.45	0.50
	Privately Funded Buffer Acres	0.77	74-79%	0.29	0.70
	Buffer Width	0.71	68-73%	0.48	0.49

2.0 Approach

Lumley (2010) proposes applying ratios or general linear models (GLMs) for adjusting survey results to account for under- or over-reporting. Because the author noted that GLMs will generally result in estimates with smaller confidence intervals, the GLM method was chosen for this technical memorandum. In this analysis, we used the R integrated suite of software facilities (R Core Team, 2016) and the “survey”[2] package (Lumley 2004 and 2016). Note that similar analytical tools are available in SAS®. Advantages of using a survey-based analytical tool over traditional GLM tools include the abilities to correctly compute the standard errors for a variety of sampling strategies and to account for finite populations.

Selection and development of a model should consider the available data. Figure 1 displays the verified acreage as a function of self-reported acreage using the PSU/DEP verification data for row crops, and is typical of data sets for other BMPs in the PSU/DEP study related to acreage estimates. Although the PC is 85 percent, 371 of the correctly classified results are attributed to observations with zero reported and zero verified acreage (green circle). In Figure 1, there are 70 errors of omission, i.e., the observations in the blue rectangle, and 35 errors of commission, i.e., the observations in the black dashed rectangle. There is one observation with a reported acreage of 11,000 that appears to be an outlier relative to the other data.

Given the characteristics of the above data set (i.e., the large number of zero reported acreage), it is recommended to develop a general linear model that accounts for the zero and non-zero reported acreage separately. This can be achieved by using the model shown in Equation 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_1 X_2 \qquad \text{Eq. 1}$$

where

Y = the verified acreage,
X_1 = indicator variable (0: reported acreage=0, 1: reported acreage>0),
X_2 = reported acreage, and
$\beta_0, \beta_1,$ and β_2 = regression coefficients.

Equation 1 can be simplified by substituting in 0 and 1 for X_1 to yield Equation 2.

$$Y = \begin{array}{cc} \beta_0 & for\ X_1 = 0 \\ (\beta_0 + \beta_1) + \beta_2 X_2 & for\ X_1 = 1 \end{array} \qquad \text{Eq. 2}$$

As can be seen from Equation 2, $\beta_0 + \beta_1$ is the y-intercept and β_2 is the slope for non-zero reported acreage observations.

[2] https://CRAN.R-project.org/package=survey

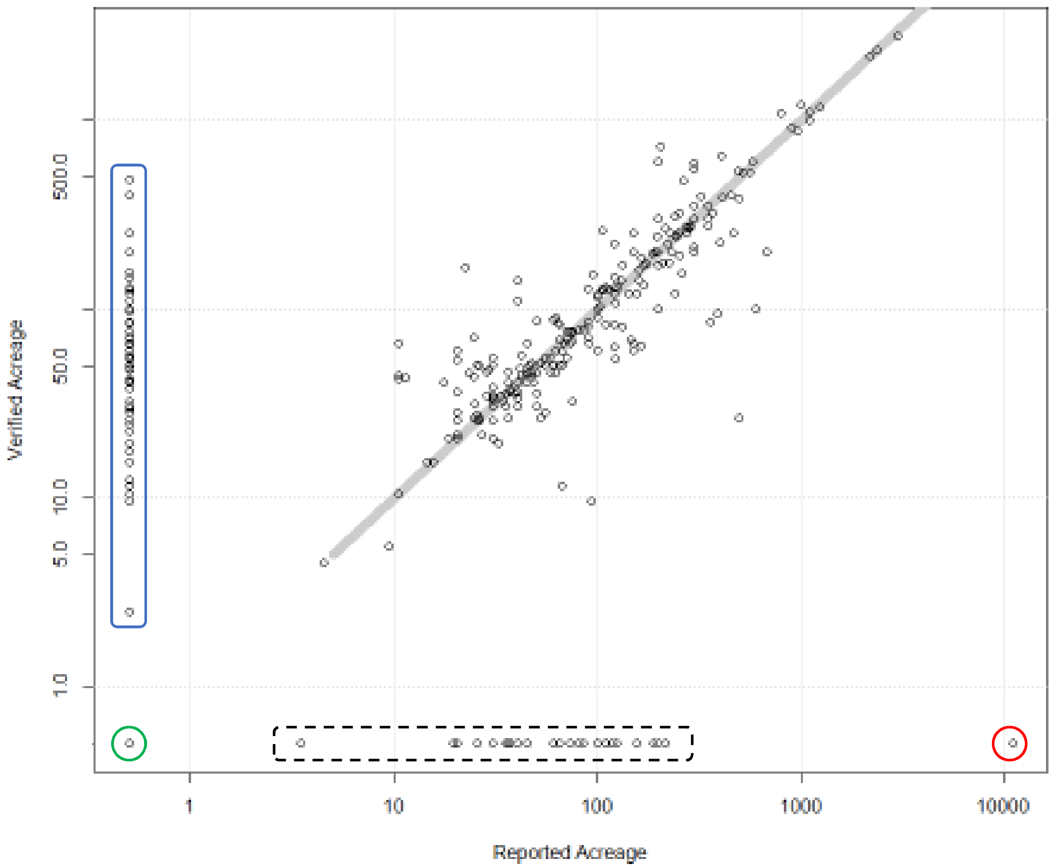

Figure 1. Verified acreage as a function of self-reported acreage for row crops. (All points are offset by adding 0.5 to facilitate plotting on a log scale. Points within blue rectangle: 70 observations with 0 reported acres and >0 verified acres. Points within black dashed rectangle: 35 observations with >0 reported acres and 0 verified acres. Points within green circle: 371 observations with 0 reported and verified acres. Points within red circle: 1 potential outlier with 11,000 reported and 0 verified acres. Grey line: 1:1 slope.)

Depending on the model fit, it may be appropriate to set the y-intercept term to zero. This can be achieved by introducing another indicator variable, X_0, which is 1 for zero reported acreage and zero otherwise (i.e., the opposite of X_1). The general equation is

$$Y = \beta_0 X_0 + \beta_2 X_1 X_2 \qquad \text{Eq. 3}$$

and the simplified model (substituting in for X_0 and X_1) is

$$Y = \begin{matrix} \beta_0 & for\ X_1 = 0\ and\ X_0 = 1 \\ \beta_2 X_2 & for\ X_1 = 1\ and\ X_0 = 0 \end{matrix} \qquad \text{Eq. 4}$$

In either case (i.e., Equation 1 or 3), the value of β_0 will correspond to the mean verified acreage for surveys where the reported acreage is zero. The functions *survey::svydesign*, *survey::svyglm*, and *stats::predict* can then be used to compute the model coefficients and estimate the state and county level totals. In our application of *survey::svydesign*, we set the strata argument to county because the procedure to select samples from the returned surveys was based on a post-stratification based on county.

3.0 Application for Statewide Estimate

Results from applying the approach described in Section 2 to the DEP/PSU row crop data are presented here. Note that it is necessary to have two or more observations per county to apply the strata argument. For the row crop data, the single samples in the verification data set for Elk and Jefferson counties were aggregated with Clearfield county; Sullivan county with Columbia county; and Wyoming County with Luzerne.

The state level results and model fits are shown in Table 2 and Figure 2, respectively. The red line uses the Equation 1 model and the entire verification data set. The blue line also uses the Equation 1 model but excludes the outlier circled in Figure 1. Finally, the black line uses the Equation 3 model and excludes the outlier.

The Equation 3 model is preferred given the lower standard error and visual inspection of Figure 2. The state estimate of 364,850 acres has 90% confidence intervals of 347,508—382,191 acres. Note, that the 90% confidence intervals do not contain the reported acreage of 335,250.

Table 2. Statewide row crops estimates.

Model	Estimated State Total	Standard Error	90% Lower Confidence Level	90% Upper Confidence Level
Equation 1	418,463	33,342	363,615	473,310
Equation 1 (exclude outlier)	355,062	15,014	330,364	379,760
Equation 3 (exclude outlier)	364,850	10,542	347,508	382,191

The approach to developing a model should generally follow the same best practices that would be used for any regression. For example, if there were enough county-level samples taken, then it might make sense to evaluate whether to add county as a covariate. Adding county to the Equation 3 model and simplifying would result in county-specific β_0 values as shown in Equation 5 where i represents the county.

$$Y = \begin{matrix} \beta_{0,i} & for\ X_1 = 0\ and\ X_0 = 1 \\ \beta_2 X_2 & for\ X_1 = 1\ and\ X_0 = 0 \end{matrix} \qquad \text{Eq. 5}$$

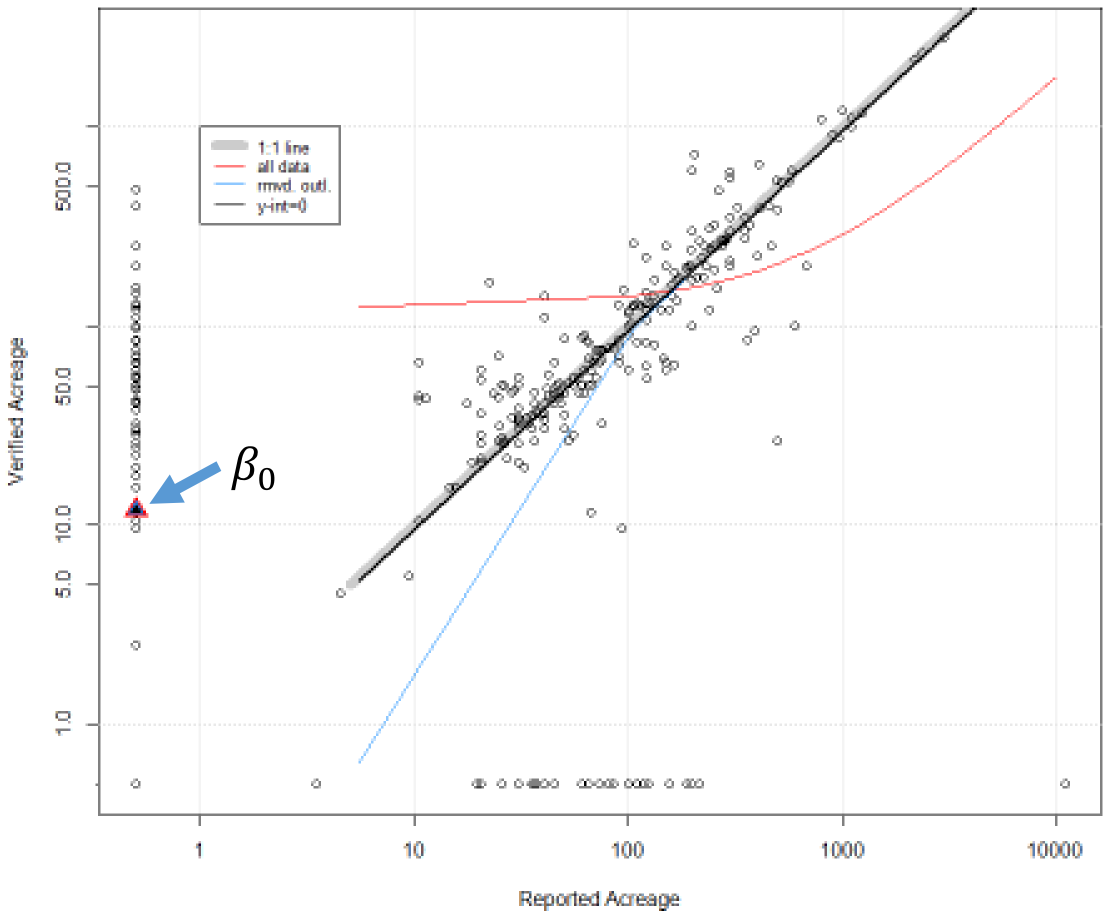

Figure 2. Verified acreage as a function of self-reported acreage for row crops together with model fits. (Fitted GLMs: Red line—Equation 1; Blue line—Equation 1, exclude outlier; Black line—Equation 3, exclude outlier.)

4.0 Hypothetical Extension to County Estimates

The Equation 3 model developed in Section 3.0 can also be applied to county level data. However, county level summary information was not available for this analysis. Therefore, a hypothetical county summary of row crop data was created for demonstration. Table 3 presents the hypothetical number of returned surveys, the number of surveys with zero reported acreage, the number of surveys with non-zero reported acreage, and total reported acreage for each county. Figure 3 presents the hypothetical predicted row crop acreage with 90% confidence intervals.

Table 3. Hypothetical county summary for row crops.

County	Returned Surveys	Surveys with Zero Reported Acreage	Surveys with Non-zero Reported Acreage	Total Reported Acreage
Adams	210	153	57	9,513
Bedford	191	153	38	2,072
Berks	96	38	58	3,952
Blair	124	86	38	5,228
Bradford	296	220	76	10,025
Cambria	57	57	-	-
Centre	229	105	124	11,050
Chester	172	86	86	6,457
Clearfield[A]	57	19	38	2,270
Clinton	67	29	38	2,113
Columbia[A]	191	162	29	1,050
Cumberland	191	124	67	20,453
Dauphin	105	38	67	11,315
Franklin	372	210	162	18,000
Fulton	105	67	38	4,227
Huntingdon	115	77	38	4,844
Indiana	38	38	-	-
Juniata	105	57	48	6,000
Lackawana	29	19	10	196
Lancaster	1,500	793	707	99,154
Lebanon	201	86	115	15,407
Luzerne[A]	76	67	9	74
Lycoming	240	173	67	5,137
McKean	38	29	9	150
Mifflin	124	57	67	5,146
Montour	115	77	38	7,726
Northumberland	124	86	38	8,750
Perry	201	115	86	15,649
Potter	67	67	-	-
Schuylkill	143	76	67	4,130
Snyder	162	143	19	9,809
Somerset	38	29	9	352
Susquehanna	267	181	86	2,369
Tioga	220	172	48	5,804
Union	143	76	67	6,700
Wayne	29	19	10	125
York	344	229	115	30,003
Total	**6,782**	**4,213**	**2,569**	**335,250**

[A] Elk and Jefferson, Sullivan, and Wyoming counties were assumed to be aggregated with Clearfield, Columbia, and Luzerne counties, respectively.

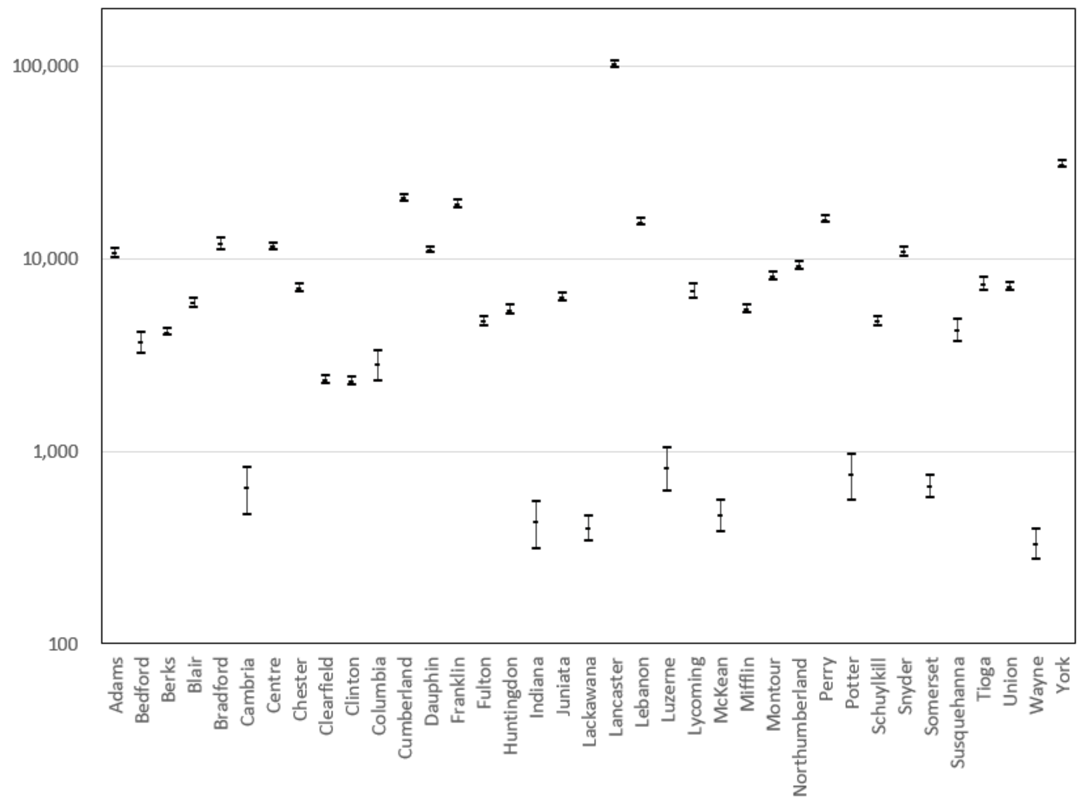

Figure 3. Hypothetical predicted row crop acreage with 90% confidence intervals.

5.0 Summary and Discussion

This technical memorandum presents an overview of a procedure that could be used to evaluate a self-certified assessment inventory (e.g., mail-in survey) that includes the follow-up independent verification using a stratified random sample of the returned mail surveys.

The general requirements for the procedure, as portrayed in this technical memorandum, include the following:

- A detailed verification data set which includes the county name, reported acreage and verified acreage. A minimum of two observations are needed in each county.
- County- and state-level summary information that includes the number of returned surveys, the number of surveys with zero reported acreage, the number of surveys with non-zero reported acreage, and total reported acreage for each county and the state overall.

With the above information, it is possible to compute overall metrics such as PC, HR, and FAR as well as state- and county-level total acreage estimates as illustrated in Table 4. The Agriculture Workgroup may want to consider both these metrics and the procedure presented here when developing criteria for determining the suitability of data collected from a producer survey.

Table 4. Summary of row crop information with 90 percent confidence intervals.

Subcategory	Reported State Acreage	Proportion Correct (PC)	Hit Rate (HR)	False Alarm Ratio (FAR)	Adjusted State Acreage
Row Crops	335,250	85% (83-87%)	77% (73-81%)	13% (10-17%)	364,850 (347,508-382,191)

6.0 References

DEP. 2016a. Pennsylvania Department of Environmental Protection QAPP addendum verification program. Pennsylvania Department of Environmental Protection, March 4, 2016, 132 p. Accessed March 6, 2017. http://www.chesapeakebay.net/documents/PA_BMP_Verification_QAPP_Final.pdf

DEP. 2016b. Presentation: Pennsylvania farm conservation practice inventory Chesapeake Bay Ag workgroup July 14, 2016. Pennsylvania Department of Environmental Protection. (2016pafarmconservationpracticeinventory.pdf)

Lumley, T. 2004. Analysis of complex survey samples. Journal of Statistical Software 9(1):1-19.

Lumley, T. 2010. Complex Surveys: A Guide to Analysis Using R. Hoboken, N.J., John Wiley.

Lumley, T. 2016. survey: analysis of complex survey samples. R package version 3.31-5.

PSU. 2016. BMP verification methodology report. Penn State University.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical. Computing, Vienna, Austria. URL https://www.R-project.org/.

Royer, M. 2016. Email communication with Matt Royer (November 8, 2016), Director of the Agriculture and Environment Center, Penn State University.

Tetra Tech. 2016. Assessment of Penn State University survey of farms in Pennsylvania. Prepared for U.S. Environmental Protection Agency, Chesapeake Bay Program Office and Pennsylvania Department of Environmental Protection, by Tetra Tech Inc., Fairfax, VA. Accessed March 6, 2017. http://www.chesapeakebay.net/channel_files/24633/tt_review_psu_survey_assessment_12-9-2016_final.pdf