Enhancing the identification of error checking in the Cow Herd Appraisal Performance Software (CHAPS) benchmarks using SAS procedures

J.M Ramsay[1], L.L. Hulsman Hanna[2] and K.A. Ringwall[1]

The Cow Herd Appraisal Performance Software (CHAPS) has been used for almost 30 years as a management tool, establishing beef production benchmarks. Our objective was to develop a process to evaluate CHAPS datasets and benchmarks for errors using SAS procedures. SAS analysis will expose dataset errors that most likely occurred during data transcription and as the result of differences in rounding methods between the SAS and CHAPS benchmarks. Identifying sources of errors in the CHAPS datasets and benchmarks will improve the CHAPS program and benchmark data.

[1]Dickinson Research Extension Center, NDSU
[2]Department of Animal Sciences, NDSU

Summary

The Cow Herd Appraisal Performance Software (CHAPS) has been used as a management tool by the beef industry since 1985. CHAPS calculates beef production benchmarks as five-year rolling averages and herd benchmarks for each year. Those benchmarks include calving distributions; reproduction traits such as pregnancy and weaning percentages; calf weight-related benchmarks such as birth weight, weaning weight, average daily gain and pounds weaned per cow exposed; and cow age, weight and condition. Overall benchmarks are calculated from the herd-year

benchmarks (five years of data from herds with a minimum of 50 cows). The benchmarks allow producers to examine numerous reproduction and production traits simultaneously to evaluate herd performance and make necessary changes. Data errors can affect data quality and accuracy; therefore, SAS procedures were used to evaluate the 2012 CHAPS benchmarks and datasets. The SAS program exposed errors in the CHAPS dataset, including inaccurate and missing data that likely was caused by data transcription. In addition, differences in rounding between SAS and CHAPS methods resulted in differences in the herd-year benchmarks, which should be considered as potential sources of error. Future evaluations of CHAPS data errors will include an assessment of error rates. SAS will continue to be used to ensure the accuracy and completeness of the CHAPS datasets and benchmarks, allowing producers to optimize herd performance through informed management decisions.

Introduction

The Cow Herd Appraisal Performance Software (CHAPS) was developed by the North Dakota State University Extension Service and North Dakota Beef Cattle Improvement Association as a management tool for beef producers. CHAPS is a data-intensive system for collecting, storing and evaluating beef production (Ringwall, 2004; Ringwall et al., 1992). The foundation of CHAPS is the benchmarks, including reproduction and production traits, calculated using equations derived from an Integrated Resource Management Standardized Performance Analysis (SPA; McGrann, 2010). CHAPS has been reviewed and approved by the National Cattlemen's Beef Association following guidelines set by the Beef Improvement Federation (Ringwall, 2004; Ringwall et al., 1992).

CHAPS measures Standardized Performance Analysis (SPA) performance traits for each herd and year to establish industry benchmarks producers can use to measure progress and identify potential areas for change. The benchmarks are calculated as five-year rolling averages from selected CHAPS herds, which must include a minimum of 50 cows.

Five-year averages buffer against rapid changes in performance traits and allow producers to gauge production against solid indicators to set and modify production goals. The overall goal of CHAPS is to ensure SPA production traits are balanced and simultaneously evaluated within a beef operation while assisting producers with total herd evaluation through managerial reports (Ringwall et al., 1992).

CHAPS was developed in 1985 in MS-DOS and updated to Windows in 2000. CHAPS provides the beef industry with reproduction and production benchmarks traits that are published on the CHAPS website (www.chaps2000.com/benchmarks.htm) and discussed in *BeefTalk* (www.beeftalk.com/), a column that Kris Ringwall writes and North Dakota State University Agricultural Communication makes available to the news media.

Reproduction benchmarks include calving distributions, pregnancy, calving and weaning percentages, culling and replacement percentages, and pounds weaned per cow exposed. Production benchmarks include birth weight and weaning weight for bulls, heifers, steers and overall; age at weaning; average daily gain; weight per day of age; frame score; cow age; cow weight at weaning; cow condition at weaning; and adjusted 205-day weaning weight, where all weights are reported in pounds.

CHAPS stores individual herd data as an MS-Access97 database separated into data tables containing individual cow-calf data according to an event in the cow-calf cycle (CHAPS, 2000). For each year, herd means for each benchmark trait are calculated using the "Herd Analysis" function of CHAPS. This function uses data that producers or NDSU data management specialists enter into CHAPS to generate a "Herd Report," This report includes reproduction and production benchmark traits (percentages and means) for each year, known as the herd-year benchmarks. NDSU data management specialists transfer the herd-year benchmarks manually into an MS-Excel 2010 (hereafter referred to as Excel) spreadsheet, identifying each CHAPS herd and year. The overall benchmarks are derived from averaging all of the CHAPS herd-year benchmarks using the average function of Excel.

Manual data transfer is prone to errors, including transcriptional errors such as transposed numbers and missing data, which can affect subsequent data analysis (Blasi et al., 2006; Hong et al., 2013). To identify potential sources of error in the CHAPS benchmarks and data, we reproduced the CHAPS herd-year and overall benchmarks for 2012 using SAS procedures and compared these with the corresponding CHAPS benchmarks. The ultimate goal of this process will be to describe the rates of error in the CHAPS datasets and benchmarks to ensure accurate and reliable CHAPS benchmarks for beef producers.

Experimental Procedures

Evaluation of the 2012 CHAPS benchmarks was performed by comparing the herd-year and overall benchmarks with those calculated by procedures using SAS software (SAS Institute Inc., Cary, N.C.). If transcription error rates were negligible, analysis using SAS was expected to yield herd-year and

overall benchmarks identical to the CHAPS analysis because they were derived from the same datasets. In the event of differences, the original Herd Reports, containing the benchmarks for each herd and year, were examined to ensure they were transcribed into the Excel spreadsheet correctly.

To calculate the benchmarks using SAS CHAPS, data tables from each herd were imported into SAS using the import procedure. For each herd, SAS data sets were created to include individual cow-calf birthdates and weights, weaning dates and weights, bull turn-out dates, culled cows and active cows. These data sets subsequently were merged and sorted to facilitate calculation of benchmark traits for each herd and year.

The herd-year reproduction benchmarks were calculated using structured query language (SQL) procedures. The herd-year production benchmarks were calculated using the means procedure of SAS. The overall 2012 benchmarks were calculated using the means procedure of SAS. The equations associated with computation of reproduction and production SPA performance measures were described previously (McGrann, 2010; Ringwall et al., 1992). The overall benchmarks, calculated by SAS, were rounded to the same decimal place as the overall benchmarks on the CHAPS website (www.chaps2000.com/benchmarks.htm).

Results and Discussion

The overall benchmarks calculated by CHAPS and SAS are described in Table 1. The CHAPS benchmarks were corrected for typographical errors but not the rounding errors, which account for the differences between the SAS and CHAPS benchmarks seen in Table 1.

The initial evaluation of the CHAPS herd-year benchmarks with SAS revealed differences between some of the CHAPS and SAS herd-year and overall benchmarks. An examination of the original CHAPS Herd Reports verified that some of the differences between SAS and CHAPS-derived benchmarks were due to transcription errors. In one example, the percentage calving at 42 days was calculated by SAS as 89 percent, whereas CHAPS reported a percentage calving at 42 days of 59 percent. When the individual CHAPS herd analysis for this herd and year was examined, the actual percentage calving at 42 days was 89 percent, indicating a transcriptional error.

Another example involved transposing two numbers in a steer

Table 1. The 2012 benchmarks, calculated as a five-year (2007-2011) rolling average of 55 Cow Herd Appraisal Performance Software herds' means and standard deviations[1]. The CHAPS values have been corrected for typographical but not rounding errors.

Benchmark Trait[2]	CHAPS[3]	SAS	n[4]
% pregnancy	93.29 ± 4.43	93.42 ± 4.87	275
% pregnancy loss	0.75 ± 1.39	0.60 ± 1.14	275
% calving	92.61 ± 4.78	92.82 ± 5.18	275
% calf death loss[5]	3.27 ± 2.81	2.90 ± 2.55	275
% weaning	90.10 ± 6.04	90.25 ± 6.38	275
% replacement	16.26 ± 10.29	15.80 ± 6.45	275
% culled	13.70 ± 5.98	13.40 ± 6.67	275
% calf death loss[6]	3.60 ± 3.08	3.15 ± 2.79	275
age at weaning (days)	191 ± 24	191 ± 24	274
% calves at 21 days	60.35 ± 21.09	61.07 ± 21.64	275
% calves at 42 days	86.19 ± 13.15	86.33 ± 14.22	275
% calves at 63 days	96.09 ± 5.10	95.62 ± 8.54	275
% calves after 63 days	3.94 ± 5.11	4.38 ± 8.54	275
steer weaning weight	574 ± 62	575 ± 63	252
heifer weaning weight	543 ± 54	545 ± 54	274
bull weaning weight	609 ± 101	608 ± 96	122
all weaning weight	564 ± 56	564 ± 56	274
pounds weaned/cow exposed	502 ± 64	501 ± 71	275
average daily gain	2.53 ± 0.27	2.53 ± 0.27	183
weight per day age	2.97 ± 0.28	2.97 ± 0.28	274
birth weight	84 ± 6	84 ± 6	194
adjusted 205-day weight	632 ± 53	633 ± 53	274
frame score	5.77 ± 0.75	5.77 ± 0.75	45
% heifers early	40.78 ± 31.28	42.07 ± 30.74	275
% heifers at 21 days	76.18 ± 23.07	76.59 ± 23.26	275
% heifers at 42 days	92.04 ± 10.81	91.82 ± 12.06	275
% cows at 21 days	57.51 ± 22.25	58.01 ± 22.83	275
% cows at 42 days	84.85 ± 14.99	85.28 ± 15.16	275
cow age (years)	5.64 ± 0.72	5.65 ± 0.73	275
cow weight at weaning	1,396 ± 128	1,393 ± 123	20
cow condition	5.74 ± 0.72	5.75 ± 0.69	29

[1]All weights and weight gains are expressed in pounds.
[2]Individual herd benchmarks for each year were calculated using CHAPS and SAS.
[3]Cow Herd Appraisal Performance Software (CHAPS).
[4]n refers to the number of herd-years used to calculate overall benchmark means.
[5]based on the number of cows exposed.
[6]based on the number of calves born.

weaning weight; what should have been recorded as 690 pounds was recorded as 960 pounds. In other cases, CHAPS benchmark data simply was missing for a particular herd and year, while SAS analysis produced a benchmark value, suggesting that the transcriber missed including this value in the Excel spreadsheet used for calculating the overall benchmarks.

CHAPS-derived benchmark data from individual herds and years were transcribed manually from Herd Reports to Excel files to calculate the five-year rolling average benchmarks. Conversely, SAS-derived benchmark data were obtained by exporting the herd benchmarks to an Excel file, eliminating transcription errors. Such errors in manually transferred data are not uncommon (Blasi, 2006; Hong et al. 2013). Future versions of CHAPS, currently under development, could interface with SAS to allow direct export of the benchmarks to other programs and more options for data analysis.

After correcting for the aforementioned errors, some of the SAS and CHAPS benchmarks still did not match. Upon further examination, we determined that differences in rounding contributed to these differences. The SAS program was coded for the values to be rounded to the same decimal place as the CHAPS website, but during manual transcription into Excel, the values often were truncated instead of rounded. These differences resulted in inconsistencies between the SAS and CHAPS herd-year benchmarks, which ultimately affected the overall benchmarks (Table 1).

Additionally, some of the benchmarks were obtained by numerous mathematical equations, which compounded the effects of the rounding differences. Inconsistencies in rounding between methods are known to affect the final results (Jeng, 2006). Future versions of CHAPS should consider rounding and its effects on the benchmark values.

Evaluating CHAPS databases ensures that accurate and complete benchmarks are available to producers. Additionally, accurate and complete data sets will be of more interest to beef researchers wishing to use CHAPS data. Data transcription and rounding errors should be considered as an important source of variation in the benchmarks. Future evaluation of the CHAPS benchmarks will include an assessment of the rates of errors in the CHAPS benchmarks and suggestions for preventing data errors.

CHAPS has provided beef producers with a method of evaluating herd performance and informing management decisions for almost 30 years. SAS evaluation of the CHAPS datasets and benchmarks further strengthens the reputation CHAPS has gained as a high-quality herd management software, which is particularly noteworthy because the CHAPS program is being rewritten to interface more effectively with current software and technologies.

Literature Cited

Blasi, D.A. 2006. Record systems for beef stocker production enterprises. Vet. Clin. N. Amer.: Food Anim. Pract. 22:475-492.

Cow Herd Appraisal Performance Software (CHAPS). 2000. North Dakota State University Extension Service. www.chaps2000.com/ (Accessed June 21, 2014)

Hong, M.K., H.H. Yao, J.S. Pedersen, J.S. Peters, A.J. Costello, D.G. Murphy, C.M. Hovens and N.M. Corcoran. 2013. Error rates in a clinical data repository: lessons from the transition to electronic data transfer - a descriptive study. BMJ Open 28:e002406.

Jeng, M. 2006. Error in statistical tests of error in statistical tests. BMC Med. Res. Meth. 6:45-47.

McGrann J.M. 2010. Standardized Performance Analysis (SPA) cow-calf reproduction, grazing, feed fed production performance measures' data organization and calculation spreadsheet. http://agrilifecdn.tamu.edu/coastalbend/files/2012/06/User-Manual-for-SPA-Reproduction-Grazing-and-SPA-Calculations-8-24-20101.pdf (Accessed June 21, 2014)

Ringwall K.A. 2004. CHAPS: basic information and decision-making power. www.beeftalk.com/archive/bt215.htm (Accessed June 21, 2014)

Ringwall, K.A., P.M. Berg and D.L. Boggs. 1992. A computer program for appraising and increasing productivity in beef cattle. Vet. Med. 87:706S-718.