Estimating in-control time for optimizing forage sampling practices

J. A. Barrientos-Blanco[1] and K. F. Reed[1,*]
[1]*Department of Animal Science, Cornell University, Ithaca, NY 14853, USA*
[*]Corresponding author: Kristan F. Reed, kfr3@cornell.edu

Abstract

Variation in forage composition significantly decreases the accuracy of diets delivered to dairy cows. Here, accuracy of the delivered diet refers to deviation from the formulated diet and can be improved by using a renewal reward model (RRM) and genetic algorithm (GA) to optimize sampling and forage composition monitoring practices for farm conditions (St-Pierre and Cobanov, 2007a). Specifically, use of quality-control-charts to monitor forage composition can identify changes in composition for which adjustment in the formulated diet will result in a better match of the nutrients delivered to cows. The objectives of this study were 1) to estimate the corn silage and alfalfa-grass haylage stable time input parameter for the RRM; and 2) optimize farm-specific sampling practices using the RRM and GA. During harvest and silo packing, we recorded the location of the harvested forage from each field-of-origin within the silos on 8 New York dairy farms. During feed-out, we collected silage samples for NIR composition analysis 3x per week for 16 weeks and recorded the field-of-origin based on the location of the forage in the silo. Silage stable time was estimated based on the field-of-origin feeding time or by using k-means clustering of forage sample composition and ranged from 2 to 9 days. Optimal sampling frequency was greater for large farms than small farms. Shorter stable time times resulted in smaller control limits than reported by St-Pierre and Cobanov (2007a) increasing the sensitivity of the quality-control-charts to identify changes nutritional composition of forage.

Keywords: forage, nutrients variability, clustering, optimize sampling practices

Introduction

Forage component variability increases the deviation of the delivered diet from the formulated diet for dairy cows. The decline in the accuracy of delivered diets increases the risk of underfeeding and overfeeding cows affecting the milk production, feed efficiency, and cow welfare. Accuracy of delivered diets can be improved by using a renewal reward model (RRM) and genetic algorithm (GA) to optimize sampling and forage composition monitoring practices for farm conditions (St-Pierre and Cobanov, 2007a). Using quality-control-charts to monitor forage composition can identify changes in compositions for which adjustment in the formulated diet will result in a better match of the nutrients delivered to cows and an economically favorable investment of farm resources (St-Pierre and Cobanov, 2007b). According to St-Pierre and Cobanov (2007b), herd size, in-control time ($1/\lambda$) of forage components, time required to receive lab results, milk production loss due to abrupt change in forage composition, and milk price are the most influential parameters of the renewal reward model to calculate the total quality cost and optimize sampling and monitoring practices. In-control time is the length of time that the forage composition is expected to remain stable, and in this study, we will refer to in-control time as the stable time ($1/\lambda$). Accurate estimates of herd size, milk price, and time required to receive laboratory results is simple. However, milk production losses related to an abrupt change in forage composition is difficult to quantify and estimating the stable time requires extensive records from frequent sampling. St-Pierre and Cobanov (2007a) used 30 days as central value (default value) for the stable time of a sensitivity analysis of the RRM. According to St-Pierre and Cobanov (2007b), 7 days is a representative stable time for the component change of forage in a tower silo with diameter of 6.1, a vertical unloading rate of 0.15 m/d, an a repose angle of 30°. The representative stable time for a horizontal silo is 28 d when the silo is filled in 30° wedges and defaced at the rate of 0.2 horizontal m/d. However, we hypothesize that the stable time of forage components ensiled in bunkers, bags, and drive over piles (DOP) is less than 30 days. According to Barrientos-Blanco et al. (2022) field-to-field variation is the most relevant source of NDF and CP variability for haylage, and of NDF and starch variability for corn silage within the farm. The feeding time of forage from a specific field-of-origin within a silo depends on the feeding rate of the individual farm and amount of forage harvested from that field but is likely less than 30 days in most cases. Thus, the high proportion of forage component variability that is attributed to variation between field suggests that the stable time of haylage and corn silage components is likely less than 30 days. In this study, we evaluated the use of the time required to feed-out forage from a field and the use of a k-means clustering analysis to estimate the average stable time of corn silage and haylage components at feed-out from 8 New York State (NYS) commercial dairy farms. The objectives of the current study were 1) to estimate the corn silage and alfalfa-grass haylage stable time input parameter for the RRM; and 2) optimize farm-specific sampling practices using the RRM and GA.

Materials and methods

Experimental data

Location of field-of-origin within a silo of corn silage and alfalfa-grass haylage were recorded at harvest during the summer of 2020 from 8 NYS dairy farms. The same silos were sampled 3x per week for 16 weeks at feed-out during the winter-spring of 2021. We linked the forage composition samples with the field-of-origin using mapped records of the silos. The method of preservation varied by farm: 3 farms preserved both forages in bunkers; 3 farms used bags for preserving both forages; 1 farm used DOP for both forages; and 1 farm used DOP for preserving corn silage and bags for preserving haylage. The maps of the locations of field-of-origins within the bunker silos were recorded by measuring the change of the height of silage on the wall sections after each field was compacted in the silo. In bags, we marked and labeled the bag with the field-of-origin ID using spray paint when the last load of the field was packed. We were not able to accurately record and create a map of field-of-origin locations for DOP farms due to the variable shape of the pile and the incoming loads from multiple fields at the same time during harvest. Further, corn silage from farm C was not included in the analysis because it was harvested later in the season to make corn snaplage. The silage maps created during harvest were used to identify the field-of-origin of the forage removed from the silos during feed-out. Measurements of the silo-face area and changes in length of the silo were used to calculate the feeding rate in meters per day and cubic meters per day at feed-out.

Determining the stable time

Stable time for corn silage and haylage was estimated using the length of the feeding time of forage from a field-of-origin and a k-means clustering approach of the forage composition during feed-out. Each approach was used to create a time-series of sequential periods of stable times of varying lengths. We then calculated the median and interquartile deviation of the stable times for corn silage and haylage of each farm as an estimate of the average stable time and its variability.

Estimating stable time with feeding time of forage from each field-of-origin method

The approach to estimate the stable time using the feeding time of forage from the fields-of-origin was implemented in two steps. First, we identified and recorded the fields-of-origin of the forage within the defaced forage feeding pile at sampling using the silo maps created during harvest. Second, we calculated the feeding time length of each field-of-origin by subtracting the initial study period day of feeding of a new field-of-origin from the last study period day of feeding forage from that field-of-origin. The feeding time of forage from each field-of-origin was assumed to be equal to sequential periods of stable time within-silo. In silo bags, the field-of-origin was identified and recorded using labeled sections of the silo bags during feed-out. In bunker silos, the feeding pile from the defaced silage contains forage from multiple fields-of-origin so the stable time was assumed to be equivalent to the time when the defaced section contained forage from the same group of fields. The mapped fields-of-origin within each silo wall section at harvest was used to identify the fields-of-origin of the silage from a feeding pile during feed-out. When feeding out from a wall section with a different list of fields-of-origin from the previous wall section began, we considered that the start of a new stable time period. We were not able to estimate the stable time of DOP farms using the feeding time for forage from each field-of-origin approach because the method used to build the silo maps did not allow us to properly identify the field-of-origin during feed-out.

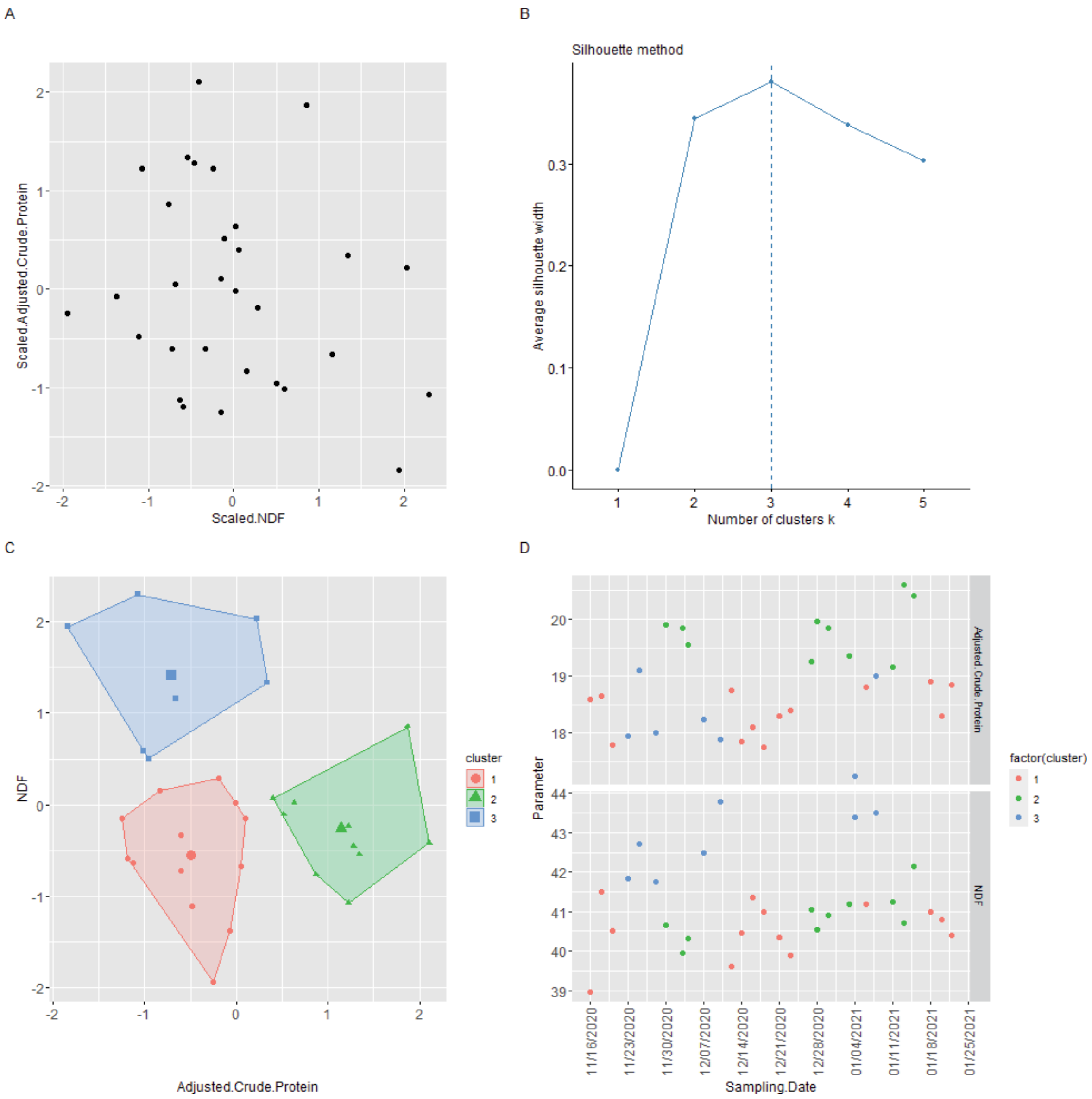

Figure 1: Scatter plot of standardized CP and NDF of haylage (A), Optimal number of clusters from silhouette method (B), clusters after k-means analysis of haylage CP and NDF (C), and time series plot of the clustered haylage CP and NDF (D).

Estimating stable time with k-means clustering

A k-means clustering method was implemented in four steps to determine the stable time for each silo of corn silage and haylage per farm. The R studio cluster package (Maechler et al., 2022) was used to perform the k-means clustering analysis in the time series dataset of corn silage and haylage components. The first step consisted of standardizing the CP and NDF of haylage and starch and NDF of corn silage ($Y = (x - \mu)/\sigma$). Second, we estimated the optimal number of clusters using the average silhouette method. Third, we clustered the standardized nutrient observations from each silo using the k-means approach. Then, for the fourth step, we calculated the median of the length of time in days within each cluster by farm and forage type. Figure 1 shows an example with the four steps of clustering to estimate the stable times of a haylage silo. In this example of haylage, we based our stable time estimates on standardized observations of CP and NDF content (Figure 1A). The silhouette analysis determined 3 to be the optimal number of clusters (Figure 1B). Then, we grouped the NDF and CP datapoints using the k-means method (Figure 1C), used the clusters to identify periods of stability, and calculated the time in days within each cluster (Figure 1D).

Optimizing sampling practices

Following the approach suggested by St-Pierre and Cobanov (2007a), we determined the optimal sample size, sampling frequency, and control limits using the renewal reward model and the genetic algorithm. The renewal reward model (RRM) is cost function used to calculate the total quality cost for control process charts developed by Lorenzen and Vance (1986). The RRM can then be used as an objective function of an iterative search technique to find the sample size, sampling frequency, and control limits at the minimum quality cost. St-Pierre and Cobanov (2007a) adapted the RRM and optimization approach for sampling and monitoring forages. The total quality cost accounts for cost related to sampling and laboratory analysis, labor cost related to monitoring forages and adjusting diets, and cost related to milk production loses when there is a change in the forage components. St-Pierre and Cobanov (2007a) suggested the use of the genetic algorithm to optimize the sample size, sampling frequency, and control limits for forages by minimizing RRM cost function. The genetic algorithm is search method for optimizing functions that uses the principles of genetic evolution that occurs in nature and is well suited for optimization problems that include a mix of discrete and continuous variables (Forrest, 1993). Sampling practices were optimized using the stable time estimates from the k-means clustering method and compared with the optimal sampling practices from a stable time fixed at 30 d. Milk prices were fixed to $0.35 per kg. Other parameters required by the RRM were obtained from the central values suggested by St-Pierre and Cobanov (2007a).

Results and discussion

The stable time estimates are reported in Table 1. and indicate that stable time for corn silage and haylage bunker, bags, and DOP might be less than the 28 d suggested by St-Pierre and Cobanov (2007b). Except for the corn silage from farm A, the estimated median stable time of corn silage ranged from 5 to 6 d using the field-of-origin feeding time method and ranged from 2 to 5 d using the k-means clustering analysis (Table 2). The feeding rate of corn silage ranged from 0.23 to 3.2 m/d. The estimated median of haylage stable time ranged from 1 to 13 d using the field-of-origin feeding time method and from 2 to 9 d using the k-means clustering approach (Table 2). Feeding rate of haylage ranged from 0.28 to 1.15 m/d. According to the sensitivity analysis conducted by St-Pierre and Cobanov (2007b), the frequency of false alarms increases dramatically as the expected stable time increases from 0 to 18 days at which point there is a large decrease in the frequency of false alarms when the stable time increases from 18 to 19 days followed by a relatively flat trend line for stable times greater than 19 days. Thus, the behavior of the system is such that the expected frequency of false alarms at stable times above 19 days is similar to the expected frequency at stable times less than 5 d. Therefore, our estimates of stable time values which are often lower than 5 d would be expected to have a similar frequency of false alarms to a system with a 28-day stable time as proposed by St-

Pierre and Cobanov (2007a) and yield a robust optimal sampling schedule and monitoring program for forages.

Table 1: Stable time (d) estimated for corn silage and haylage using field-of-origin feeding time and k-means clustering methods and feeding rate in m/d and m^3/d.

				*Feeding rate		†Stable time (d)	
Feed	Farm	Herd size	Ensiling method	m/d	m3/d	Field-of-origin feeding time	K-means clustering
Corn Silage	A	115	Bag	0.92 ± 0.39	42.81 ± 30.85	18 ± 18	32 ± 11
	C	517	Bunker	0.39 ± 0.19	230.96 ± 188.95	6 ± 8	4 ± 3
	D	552	Bunker	0.36 ± 0.14	34.44 ± 13.05	30 ± 22	5 ± 2
	E	656	Bunker	0.23 ±	1569.08 ± 0.00	10 ± 10	5 ± 4
	F	1229	Bag	3.2 ± 2.1	127.72 ± 70.5	6 ± 9	2 ± 1
	G	2064	DOP	0.68 ± 0.41	1631.39 ± 1494.07		4 ± 7
	H	3441	DOP	0.23 ± 0.14	830.51 ± 376.47		2 ± 2
Haylage	A	115	Bag	0.74 ± 0.68	40.71 ± 25.95	6 ± 4	9 ± 11
	B	309	Bag	1.15 ± 0.15	56.71 ± 17.14	1 ± 2	5 ± 9
	C	517	Bunker	0.28 ± 0.02	118.03 ± 38.01	16 ± 3	2 ± 1
	D	552	Bunker	0.29 ± 0.13	35.3 ± 17.7	1 ± 0	2 ± 1
	E	656	Bunker	0.32 ± 0.11	141.38 ± 51.74	8 ± 5	7 ± 4
	F	1229	Bag	1.15 ± 0.56	61.53 ± 114.34	4 ± 3	4 ± 4
	G	2064	Bag	0.83 ± 0.44	26.64 ± 12.32	4 ± 5	3 ± 3
	H	3441	DOP	0.37 ± 0.32	178.63 ± 122.85		2 ± 3

*Average ± standard deviation, †Median ± interquartile

Table 2 shows the optimal sampling practices determined using our stable time estimates from the k-means clustering approach for corn silage and haylage in comparison to the optimized practices using a stable time of 30 d for a theoretical standard feed proposed by St-Pierre and Cobanov (2007a). As expected, the decrease in stable times resulted in a decrease in the optimal sample size from 2 to 1 in most cases. Consistent with St-Pierre and Cobanov (2007a), the optimal sampling interval (the inverse of sampling frequency) in days decreased with the increase of farm size. The optimal control limits were lower for the shorter stable times estimated with the k-means approach than the optimal control limits for a stable time of 30 d. Lower control limits may not only increase the probability of detecting true changes in corn silage and haylage components but will likely increase the probability of false alarms. However, if the stable time is less than 5 d, we can expect similar frequency of false alarms as for systems with a stable time 19 d or greater as reported by St-Pierre and Cobanov (2007b). Therefore, we believe the optimal sampling practices determined our estimates for stable times are sensitive enough to detect small but important abrupt changes in corn silages and haylage components and also robust to the problem of excessive false alarms at farms. The lower optimal control limits for systems with stable times estimated with the k-means clustering method explains the increase of total quality costs (Table 2). Higher probability of finding abrupt changes in the forages results in an expected increased frequency of detected changes and thus increases the costs required to confirm the change of nutrient composition and to adjust the diets.

Table 2: Optimal sample size, sampling frequency, and control limit factor estimated using 30 days and farm specific stable time estimated with k-means clustering analysis for multiple size NY dairy farms.

Forage	Farm	Herd size	Stable time (d)	Sampling frequency (days)	Sample size (n)	Control limits factor	Total quality cost ($)
Corn Silage	A	115	32	12	2	1.14	$53.03
	C	517	4	4	1	0.47	$339.69
	D	552	4	4	1	0.45	$360.60
	E	656	5	3	1	0.67	$401.22
	F	1229	2	2	1	0.53	$856.68
	G	2064	4	2	2	0.82	$1,228.60
	H	3441	2	1	1	0.70	$2,303.17
Haylage	A	115	9	11	1	0.55	$77.63
	B	309	5	5	1	0.58	$204.90
	C	517	2	4	1	0.30	$380.21
	D	552	2	4	1	0.29	$385.64
	E	656	7	3	2	1.09	$367.67
	F	1229	4	2	1	0.68	$753.53
	G	2064	3	2	1	0.41	$1,309.34
	H	3441	2	1	1	0.70	$2,303.17
Standard	A	115	30	12	2	1.13	$54.19
	B	309	30	6	2	1.24	$119.65
	C	517	30	5	2	1.18	$185.06
	D	552	30	5	2	1.15	$195.91
	E	656	30	4	2	1.25	$227.45
	F	1229	30	3	2	1.23	$397.37
	G	2064	30	2	2	1.33	$637.98
	H	3441	30	2	3	1.40	$1,026.52

The time required to feed-out forage from each field-of-origin is a method that can be used to the estimate the stable time of corn silage and haylage nutrient composition. However, mapping the field-of-origin within silos is labor intensive and might not be practical for dairy farm. In addition, we are assuming a true change in the composition every time a new field-of-origin is defaced from the silo and fed to cows and this assumption might not always apply. In contrast, k-means clustering is an approach that can be implemented for a period of time at feed-out to estimate the stable time without the extra labor and costs required to map the fields within every silo.

Conclusions

According to our analysis, stable time for horizontal silos of corn silage and haylage is shorter than 28 days. K-means clustering is a method that can be implemented to estimate stable time of corn silage and haylage nutrient composition using laboratory analyses of feed-out samples.

Acknowledgments

This work is supported by the USDA National Institute of Food and Agriculture Hatch and Smith-Lever Funding under project #2019-20-130.

References

Barrientos-Blanco, J.A., Lawrence, J., and Reed, K.F. (2022) Managing diet nutrient variability through improved forage sampling practices. *The Manager, PRODAIRY. Progressive Dairy.*

Forrest, S. (1993) Genetic algorithms: principles of natural selection applied to computation. *Science* 261(5123), 872-878.

Lorenzen, T.J., and Vance, L.C. (1986) The Economic Design of Control Charts: A Unified Approach. *Technometrics* 28(1), 3-10.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2022) *Cluster Analysis Basics and Extensions.*

St-Pierre, N.R., and Cobanov, B. (2007a) A Model to Determine the Optimal Sampling Schedule of Diet Components1. *Journal of Dairy Science* 90(12), 5383-5394.

St-Pierre, N.R., and Cobanov, B. (2007b) Optimal Sampling Schedule of Diet Components: Model Robustness to Departure from Assumptions. *Journal of Dairy Science* 90(12), 5395-5404.