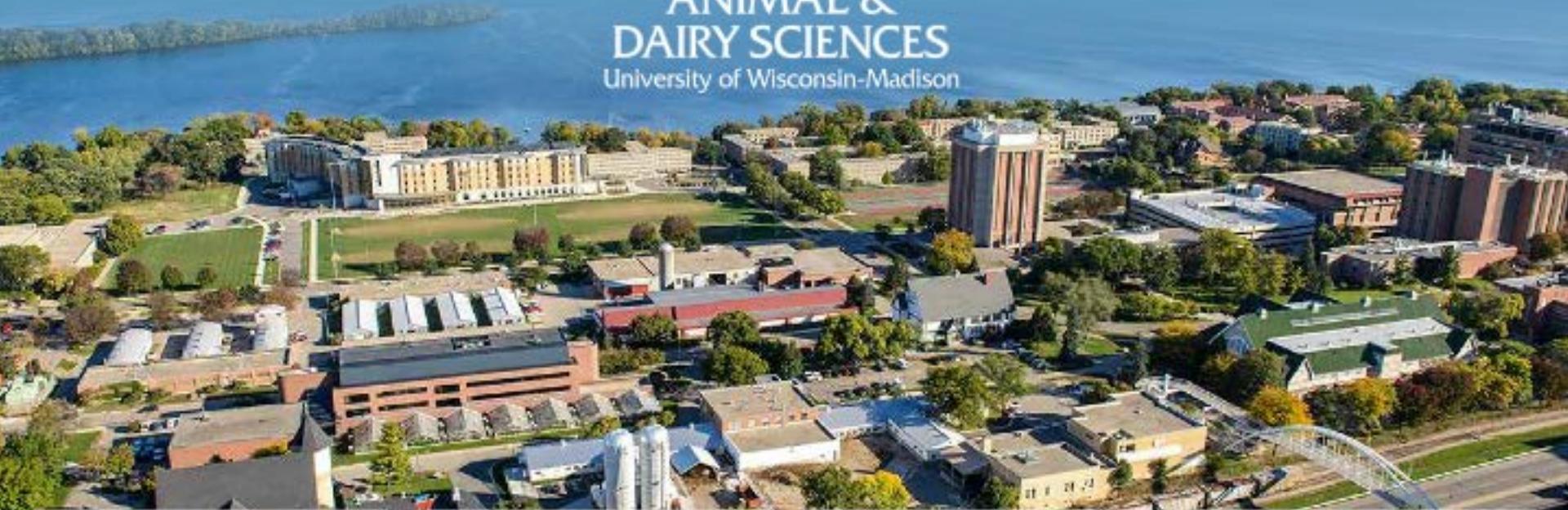




ANIMAL & DAIRY SCIENCES

University of Wisconsin-Madison



Predicting sperm production of young dairy bulls using collection history and management factors

Allison Quick¹, James Meronek², Kyle Amburn², Kevin Rozeboom², and Kent Weigel¹

1 Department of Animal & Dairy Sciences, University of Wisconsin-Madison, Madison, WI

2 ABS Global Inc., DeForest, WI

Objectives

- ❖ Develop prediction models for daily, weekly, and monthly total sperm (TS) production from collection history, health status, and management factors and compare their performance
- ❖ Assess the ability of these prediction models to forecast future TS production, as well as differences in prediction accuracy by seasonality or age of bull

- ❖ Data from a commercial AI company from 2015 to 2019
- ❖ Data aggregated into daily, weekly, and monthly TS production for HO, JE bulls 10 to 28 months of age
- ❖ Prediction Methods:
 - Linear Regression (LM)
 - Bayesian Regularized Neural Networks (BRNN)
 - Random Forest (RF)
 - Model Tree (M5P)
 - Extreme Learning Machine (ELM)
 - Multi-Layer Perceptron (MLP)
- ❖ 5-fold cross validation of training sets
 - R 3.5.1 and caret package

Materials and Methods

	Datasets		
	Daily	Weekly	Monthly
Records (n)	43,918	23,404	5,127
Holstein	39,172	20,661	4,441
Jersey	4,746	2,743	686
Bulls (n)	1,037	1,003	664
Holstein	900	872	570
Jersey	137	131	94

- ❖ Potential Explanatory Variables:
 - Age at collection, first successful collection, and arrival at AI stud
 - Collection frequency
 - Collection interval (days since previous collection)
 - Breed of bull
 - Barn where housed and collected
 - Year-season of collection
 - Scrotal circumference (SC) at 10 to 11 months of age
 - Health events
 - TS on the three most recent collection dates (lags)

Results: Summary Statistics

	Min	Mean	Max
Total sperm per Week (billion cells)	0.1	15.1	82.2
Ejaculates per Week	1	3.5	10
Total Sperm per Month (billion cells)	0.1	49.9	252.5
Collections per Month	1	5.9	15
Arrival Age (mo)	1	6.2	15
Collection Age (mo)	10	17.7	28

Results: Model Selection

	Accuracy of Total Sperm Predictions					
	Daily		Weekly		Monthly	
Method	RMSE	r	RMSE	r	RMSE	r
LM	3.3	0.73	5.1	0.82	16.9	0.87
RF	3.0	0.77	4.7	0.84	14.5	0.90
BRNN	3.0	0.77	4.8	0.84	15.6	0.91
M5P	3.1	0.76	4.9	0.84	15.7	0.91
MLP	3.6	0.65	7.7	0.73	28.7	0.69
ELM	3.4	0.71	5.7	0.78	18.2	0.84

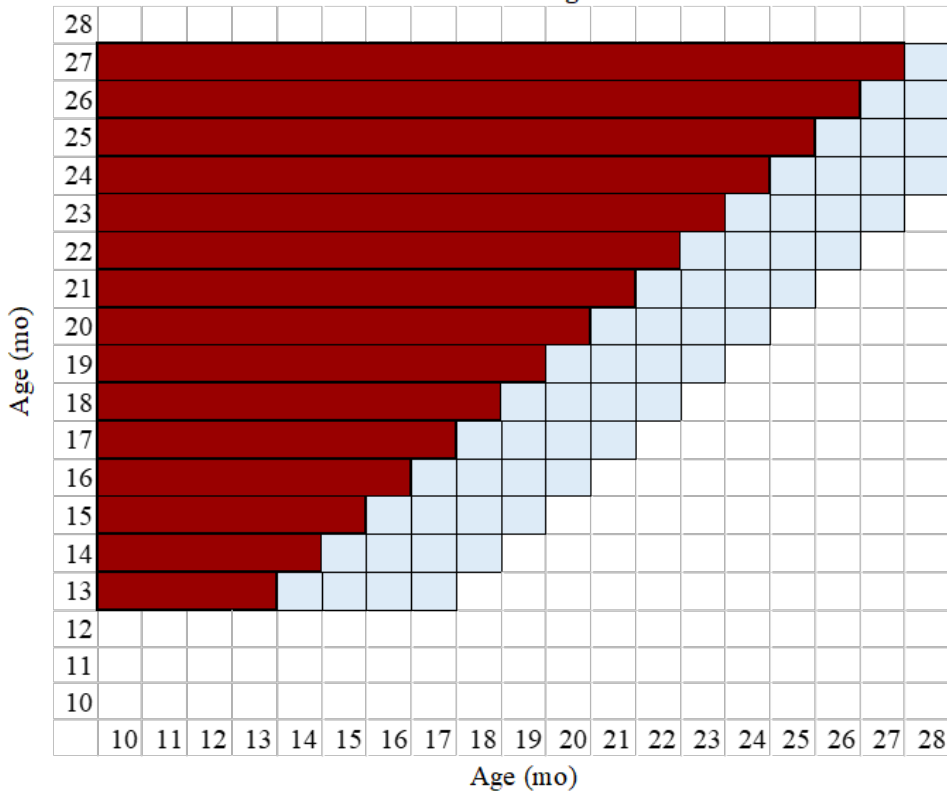
Methods: Age-Based Prediction

Training

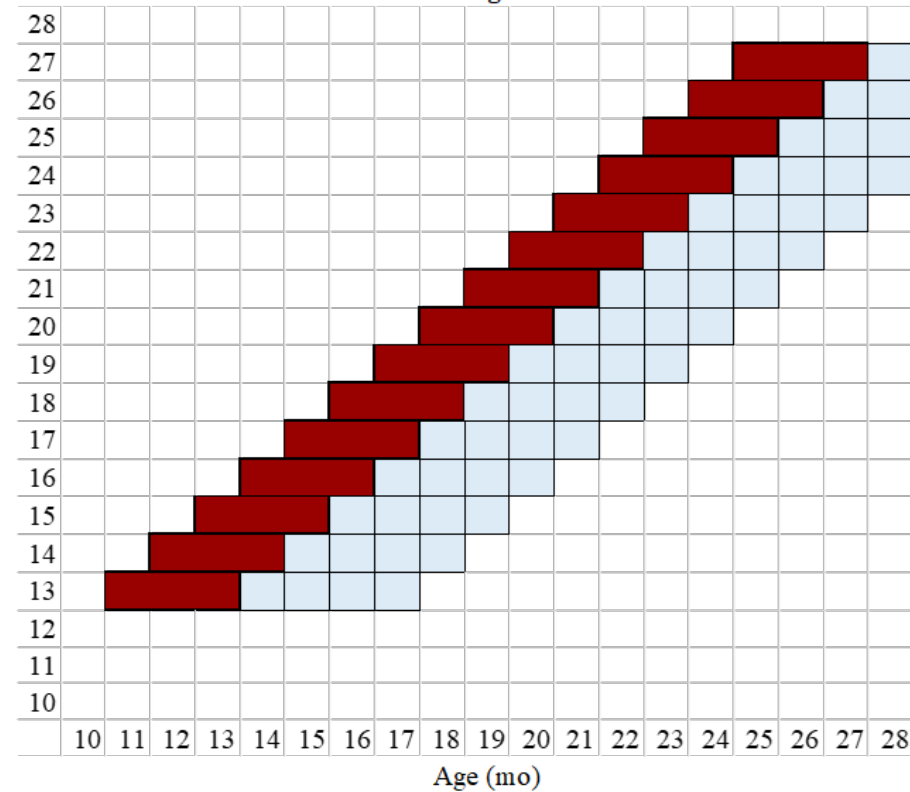
Testing

Training sets split by age

Additive Training Window



Fixed Training Window = 3



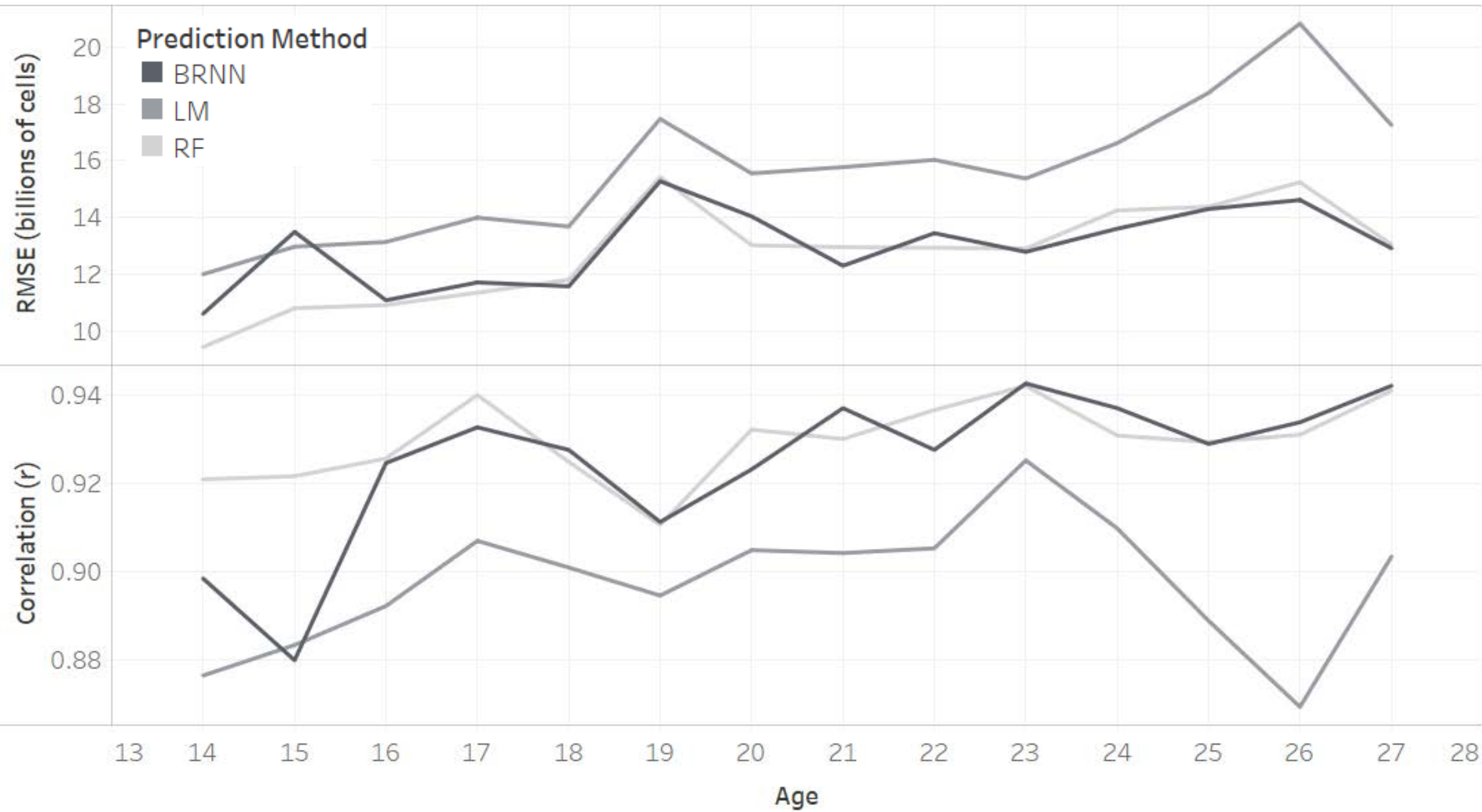
Monthly

- 2017-2019, rolling months
- Training sets:
 - Additive
 - Fixed:
 - 1 year, 4 month
- Testing sets:
 - Subsequent 4 months
- LM, RF

Weekly

- 2017-2019, rolling weeks
- Training sets:
 - Additive (LM only)
 - Fixed:
 - 1 year, 1 month
- Testing sets:
 - Subsequent 4 weeks
- LM, RF

Results: Age-Based Prediction



Results: Age-Based Prediction

Model	Training Set	Age + 1 month		Age + 2 months		Age + 3 months		Age + 4 months	
		RMSE (SD)	r (SD)	RMSE (SD)	r (SD)	RMSE (SD)	r (SD)	RMSE (SD)	r (SD)
LM	Additive	15.7 (2.4)	0.90 (0.02)	16.4 (2.0)	0.89 (0.02)	17.0 (1.8)	0.89 (0.02)	17.7 (1.6)	0.89 (0.02)
RF	Additive	12.7 (1.8)	0.93 (0.01)	13.4 (1.4)	0.93 (0.01)	14.0 (1.3)	0.92 (0.01)	14.5 (1.2)	0.92 (0.02)
BRNN	Additive	13.0 (1.4)	0.93 (0.02)	13.7 (1.0)	0.92 (0.02)	14.6 (1.6)	0.92 (0.03)	15.3 (1.7)	0.91 (0.03)
LM	Fixed 3 mo	15.1 (2.0)	0.90 (0.01)	15.7 (1.6)	0.90 (0.02)	16.3 (1.5)	0.90 (0.02)	17.0 (1.3)	0.89 (0.02)
RF	Fixed 3 mo	12.9 (1.8)	0.93 (0.01)	13.5 (1.5)	0.93 (0.01)	14.2 (1.4)	0.92 (0.01)	14.7 (1.2)	0.92 (0.01)
BRNN	Fixed 3 mo	13.5 (1.9)	0.92 (0.02)	14.5 (1.9)	0.91 (0.02)	15.5 (2.4)	0.91 (0.03)	16.3 (3.4)	0.91 (0.03)

Results: Date-Based Monthly



Results: Date-Based Monthly

Model	Training Set	Date + 1 month		Date + 2 months		Date + 3 months		Date + 4 months	
		RMSE (SD)	r (SD)	RMSE (SD)	r (SD)	RMSE (SD)	r (SD)	RMSE (SD)	r (SD)
LM	Additive	15.3 (2.2)	0.90 (0.02)	15.4 (2.0)	0.90 (0.02)	15.5 (2.1)	0.90 (0.02)	15.6 (2.1)	0.90 (0.02)
RF	Additive	11.9 (2.4)	0.94 (0.02)	12.0 (2.5)	0.94 (0.02)	11.8 (2.3)	0.94 (0.02)	11.8 (2.4)	0.94 (0.02)
LM	Fixed 1 yr	15.3 (2.2)	0.90 (0.02)	15.6 (2.0)	0.90 (0.02)	15.8 (2.1)	0.90 (0.02)	15.9 (2.0)	0.89 (0.02)
RF	Fixed 1 yr	12.1 (2.5)	0.94 (0.02)	12.2 (2.5)	0.94 (0.02)	11.9 (2.2)	0.94 (0.02)	11.9 (2.3)	0.94 (0.02)
LM	Fixed 4 mo	15.6 (2.5)	0.90 (0.03)	15.8 (2.6)	0.89 (0.02)	15.9 (2.6)	0.89 (0.02)	16.3 (2.9)	0.89 (0.03)
RF	Fixed 4 mo	13.1 (2.2)	0.93 (0.02)	13.1 (2.0)	0.93 (0.02)	13.0 (1.9)	0.93 (0.01)	13.1 (1.7)	0.93 (0.01)

Results: Date-Based Weekly

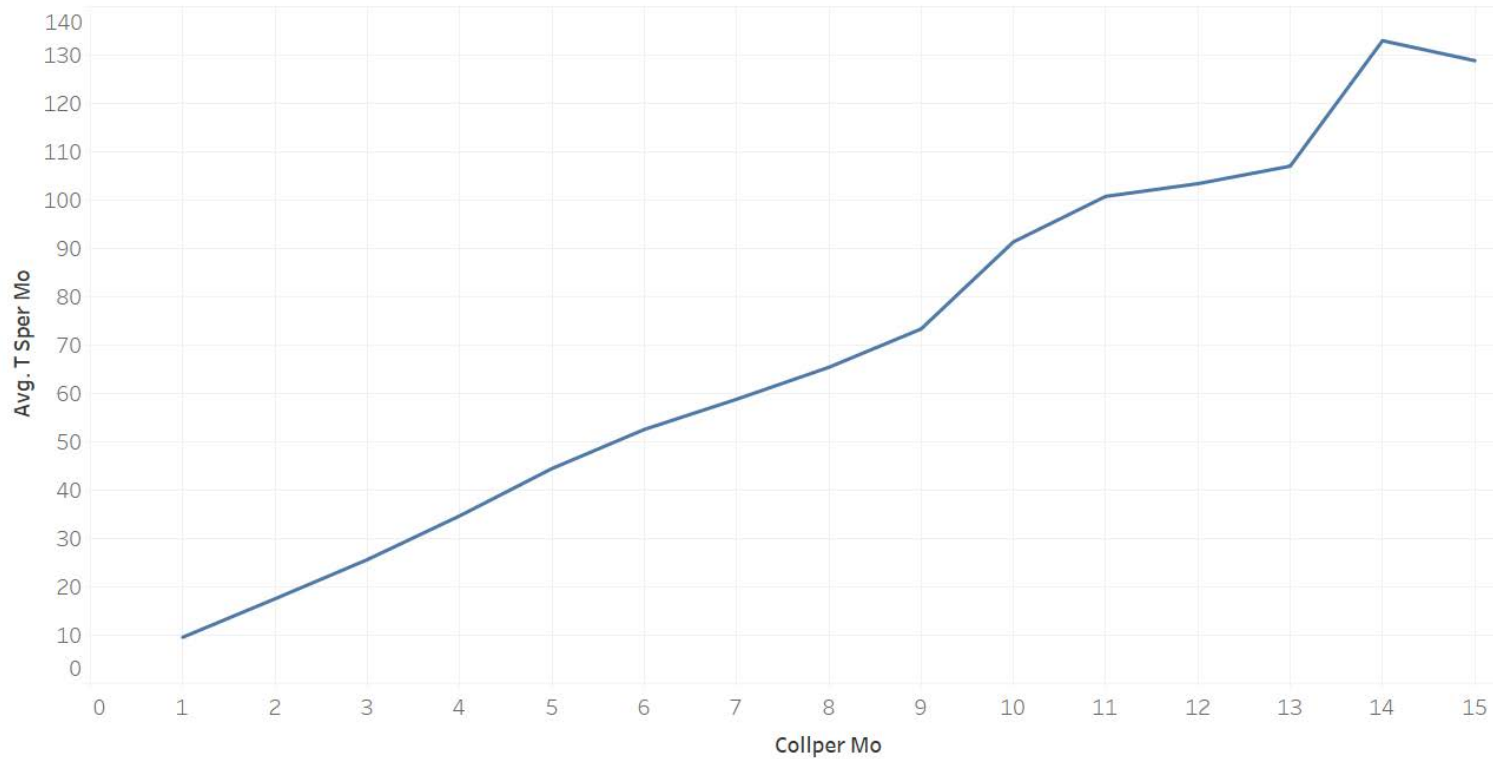
Model	Training Set	Date + 1 week		Date + 2 weeks		Date + 3 weeks		Date + 4 weeks	
		RMSE (SD)	r (SD)	RMSE (SD)	r (SD)	RMSE (SD)	r (SD)	RMSE (SD)	r (SD)
LM	Additive	4.8 (0.6)	0.84 (0.04)	4.8 (0.6)	0.84 (0.04)	4.8 (0.6)	0.84 (0.04)	4.8 (0.6)	0.84 (0.04)
LM	Fixed 1 yr	4.8 (0.6)	0.84 (0.04)	4.8 (0.6)	0.84 (0.04)	4.8 (0.6)	0.84 (0.04)	4.8 (0.6)	0.84 (0.04)
RF	Fixed 1 yr	4.4 (0.6)	0.86 (0.04)	4.4 (0.6)	0.86 (0.04)	4.4 (0.6)	0.86 (0.04)	4.5 (0.6)	0.86 (0.04)
LM	Fixed 1 mo	4.9 (0.6)	0.84 (0.04)	4.9 (0.6)	0.84 (0.04)	4.9 (0.6)	0.84 (0.04)	4.9 (0.7)	0.83 (0.04)
RF	Fixed 1 mo	4.7 (0.6)	0.85 (0.04)	4.7 (0.6)	0.85 (0.04)	4.7 (0.6)	0.85 (0.04)	4.7 (0.6)	0.85 (0.04)

Results: Important Variables

Collection Frequency

- Correlation with monthly TS: 0.77

Coll Frequency vs TS per mo



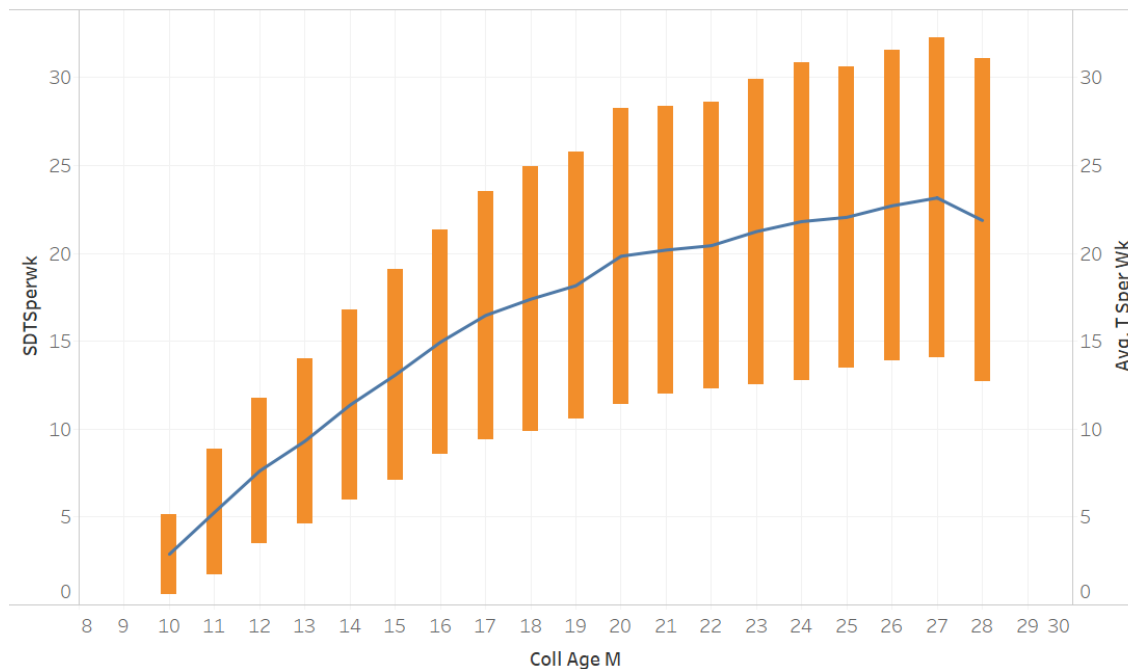
Results: Important Variables

Standardized TS at previous collections

Correlations	Lag 1	Lag 2	Lag 3
TS per month	0.386	0.413	0.428
TS per week	0.555	0.575	0.581

Collection age

Ave and SD TS by age



Conclusions

- ❖ TS production can be predicted with >90% accuracy when using RF or BRNN
- ❖ Fixed-length training sets can improve computational feasibility without great loss of predictive ability
- ❖ Most important predictors:
 - Collection frequency
 - Standardized TS at previous collections
 - Collection age
- ❖ In practice, RF model can be used to compute weekly and monthly TS predictions
- ❖ Future studies should improve prediction accuracy and develop decision support tools for bull management

Acknowledgement

ABS Global Inc.



- Amann, R. P. 1976. Bull management to maximize semen output. Page 1 in Proc. 6th NAAB Tech. Conf., February 20 to 21, 1976.
- Everett, R.W., B. Bean, and R.H. Foote. 1978. Sources of Variation of Semen Output. *J. Dairy Sci.* 61:90–95. doi:10.3168/jds.s0022-0302(78)83555-3.
- Mathevon, M., M.M. Buhr, and J.C.M. Dekkers. 1998. Environmental, Management, and Genetic Factors Affecting Semen Production in Holstein Bulls. *J. Dairy Sci.* 81:3321-3330.
- Murphy, E.M., A.K. Kelly, C. O’Meara, B. Eivers, P. Lonergan, and S. Fair. 2018. Influence of bull age, ejaculate number, and season of collection on semen production and sperm motility parameters in holstein friesland bulls in a commercial artificial insemination centre. *J. Anim. Sci.* 96:2408–2418. doi:10.1093/jas/sky130.
- Taylor, J.F., B. Bean, L.C.E. Marshall, and J.J. Sullivan. 1985. Genetic and Environmental Components of Semen Production Traits of Artificial Insemination Holstein Bulls. *J. Dairy Sci.* 68:2703–2722. doi:10.3168/jds.S0022-0302(85)81155-3.