320 Automated collection and processing of data in livestock farms. J. Koltes*, Iowa State University, Ames, IA.

Smart livestock farming will enable real-time animal management to maximize health and efficiency. The dairy industry has been generating “big data” for many years for use in genetic and management decision making. Thus, a large existing data chain exists to implement smart farming “big data” for many years for use in genetic and management decision making. The dairy industry has been generating big data animal science. For producers, affordability, accuracy and sensitivity of analytic tools will determine if efforts to create automated collection and processing systems will be fruitful. Challenges in creating these systems include: network connectivity, sensor range, data dimensionality, data ownership/ privacy concerns, and development of effective predictive analytic cyberinfrastructure. Processing (i.e., cleaning, denoising, normalizing) should ideally allow rapid data mining, analytics and integration. Given the high value of data, preservation (i.e., backup) will be critical to prevent information loss. Capturing provenance (e.g., edits/ updates) and metadata will be important for downstream analytics. Livestock data has a broad array of uses from on-farm decision making to research, thus a broad range of animal and computational scientists should be involved in developing informatics systems. In principle, all data would be FAIR (Findable, Accessible, Interoperable, and Reusable) to encourage new research and expanded application. Protecting data ownership and maintaining privacy is an important consideration as downstream processing will likely require centralized cloud computing resources. Development of databases and software that facilitate machine learning, artificial intelligence or other prediction methods would help in improving computational efficiency and reduce wait time for analytic information. Thoughtful development of cyber-physical systems for smart farming could facilitate a new era of big data animal science. For producers, affordability, accuracy and sensitivity of analytic tools will determine if efforts to create automated collection and processing systems will be fruitful.

Key Words: cyberinfrastructure, precision livestock farming, sensors

321 Use of big data to monitor herd health. H. White*, University of Wisconsin Madison, Madison, WI.

The ability of early detection and treatment to ameliorate negative outcomes associated with health indices have emphasized the value of early detection and treatment. Although this is recognized, the labor-intensiveness and costs associated with monitoring herd health are cumbersome for many farms. Increasing data streams available to farms presents opportunities to use data to better monitor cow and herd health. Within the transition to lactation period, data are presented in the form of milk production and composition, fourier-transform infrared spectroscopy (FTIR) analysis, cow management records, and genomics: all components are employed to monitor postpartum onset of hyperketonemia (HYK). First attempts to predict postpartum HYK from test-day milk and performance variables incorporated into multiple linear regression models had sufficient accuracy to monitor monthly herd prevalence; however, the sensitivity (SEN) and specificity (SPC) lacked for individual cow diagnostics. Subsequent artificial neural network prediction models employing FTIR data and milk composition variables achieved 83 and 81% SEN and SPC. While this fails to reach the diagnostic goals of 90%, it is achieved without individual cow blood samples, which may justify acceptance of lower performance. The caveat is that these models depend on milk analysis which is traditionally done every 4 weeks, only catching half of the fresh cows once. Benefits to farms are greatly improved if postpartum cows are tested weekly. This also allows for close monitoring of SCC and may open the door for other herd health monitoring tools to be employed. Future improvements in these models may be achievable by tailoring models to reflect risk factors for each farm or by maximizing SEN at the expense of SPC, especially in HYK and other disorders when the cost of treatment is less than mistreating. Incorporation of genome-wide association studies identifying SNPs associated with risk for metabolic disorders may allow for at-risk cows to be identified and flagged with lower thresholds on prediction models. Alignment of data sources, and integration of existing models, may provide valuable real-time management tools to producers.

Key Words: ketosis, hyperketonemia, transition cow

322 Real-time continuous decision-making using big data. V. Cabrera*, J. Barrientos, L. Fadul, and H. Delgado, University of Wisconsin-Madison, Madison, WI.

We are developing a real-time, data-integrated, data-driven, continuous decision-making engine, a Dairy Brain. We illustrate the concept with 3 practical applications. (1) Precision feeding: all lactating cows (N=1,898 ± 115) in a farm are housed in 14 same-sized pens and fed 7 diets. Cows are allocated to each pen only according to DIM and lactation number. Every week employees move cows in and out of pens (N=256) following lists manager generates from printouts of disjointed data. A continuous, more systematic cows’ allocation to pens according to nutritional requirements is possible by automatically aggregating data streams from management, feed, DHI, and parlor records. Algorithms can provide more precise diets to more homogeneous clusters of cows allocated to pens. Preliminary results indicate the farm could increase net margin by $67,834/yr when providing pen-based more precise diets. (2) Early risk of clinical mastitis (CM): first-lactation cows under risk of developing CM can be identified by analyzing integrated data from their genomic traits scores such as Genomic Total Performance Index (GTPI), management, and laboratory results. The risk can be reassessed at every DHI report or GTPI score variation. This integrated continuous algorithm allows to identify high-risk CM cows even when their SCC is low (<200,000 cells/mL); for example, cows with GTPI > 2,000 report the lowest SCC during the entire lactation. However, animals within this category and SCC > 115,000 cells/mL should be carefully monitored because they might already be infected. (3) Predicting CM: Cows at risk of CM show higher milk conductivity and milk production than healthy cows. Continuous, integrated data could be used to permanently predict the onset of CM. The difference on milk production and conductivity between milkings analyzed with gradient boosting machine learning algorithms can predict CM 5 milkings earlier with a sensitivity of 0.82, a specificity of 0.58, and a prediction accuracy of 0.81. Farmers can count with a monitoring tool to flag cows at risk of contracting CM and follow up them closely in a preventive way. More data for training the machine learning algorithms and integration of other data streams such as genetics, sensors, and diet changes, will improve the prediction accuracy.

Key Words: decision support tools, simulation, optimization
Optimized decisions using big data analytics in dairy farms. M. Ferris*, A. Christensen, and S. Wangen, University of Wisconsin, Madison, WI.

Management decisions can be informed by real-time data streams to not only improve the economics of the farm but also positively benefit the overall health of a dairy herd or the larger environment. Decision support tools can provide data management services and analytics to exploit data streams from farm and other economic, health and agricultural sources. We will describe a decision support tool that couples an optimization model to underlying cow, herd and economic data with an easy to interpret user interface. This interface allows the user to operate the optimization model and understand the impacts of different decisions. Specific examples related to ketosis and to culling will be described, along with some discussion of security of information and control of uncertainties in mathematical modeling approaches. The proposed analytical capabilities will utilize several tools from the data mining, simulation, machine learning, and optimization disciplines. Broadly, these tools will form the basis of the Dairy Brain decision support system, which will provide management suggestions to farmers that might affect a single animal or they could broadly affect a farm’s overall business portfolio (i.e., large capital expenditures, outsourcing opportunities, and interactions with entities that have regulatory control). All of the analytical tools are operationalized in a way that allows them to accept near real-time data updates from the farm to our agricultural data hub that gathers and disseminates multiple data streams relevant to dairy operations.

Key Words: optimization, data science, decision support