Event	Montreal (CA), ICAR 2022 Annual Conference
Subject	Abstract presented manuscript as ORAL presentation

Title of the presentation

A machine learning approach to describe the welfare status of dairy herds and analyse its association with performance and economics

Presenter: Gabriel Dallago, McGill University, Canada
             E-mail: gabriel.dallago@mail.mcgill.ca

Session: Artificial intelligence for adding value (to be confirmed)

Authors: Gabriel Dallago Daniel Warner Kevin Wade Roger Cue René Lacroix Abdoulaye Baniré Diallo Elsa Vasseur

Title of the presentation: A machine learning approach to describe the welfare status of dairy herds and analyse its association with performance and economics

ABSTRACT

The implementation of animal welfare (AW) solutions is influenced by their economic impact, but the relationship between welfare and herd performance and economics is unclear. Our objective was to describe the AW status of Quebec, Canada dairy herds and analyse its relationship with herd performance and economics. The following five AW outcomes, collected at herd-level between 2016 and 2019, were extracted from the proAction® database from 2,980 herds: prevalence of lameness, hock, neck, and knee lesions and animals with a body condition score (BCS) ≤ 2. Lactanet provided cumulative energy corrected milk (ECM; kg), cumulative milk value (CAD$), length of productive life (year), animals on third or greater lactation (3+ lact; %), and the herd status index (i.e., welfare and health remote monitoring tool; HSI) on those herds. Farms were grouped based on welfare outcomes, HSI, season and barn type using self-organizing map and cluster analysis. Machine learning models were trained to predict clusters based on welfare outcomes. The best model was used to describe the clusters with partial dependency plots. Machine learning models were also trained to analyse the relationship between clusters and both performance and economic metrics by using accumulated local effect plots. Models were trained with 10-fold cross-validation on 80% of the data and evaluated on the remaining 20%. Weighted classification, synthetic minority over-sampling technique and up-sampling were also tested for the analyse of performance and economic metrics. Partitioning around medoids with three clusters had the best stability and quality. The models created with the algorithms support vector machine and weighted extreme gradient boosting machine had the highest accuracy in predicting cluster labels based on AW outcomes (0.85; 95%CI: 0.81 – 0.87) and both performance and economic metrics (0.44; 95%CI: 0.40 – 0.48), respectively. Cluster 1 (N = 717 herds) was characterized as having a high prevalence of lameness, hock lesions and animals with BCS ≤ 2, and not having the highest ECM. Cluster 2 (N = 604 herds) had a high prevalence of both neck and knee lesions, and it was associated with lower ECM, lower 3+lact, and high milk value. Lastly, cluster 3 (N = 1,659 herds) was associated with a low prevalence of all AW issues as well as high ECM, high 3+lact and lower milk value. By using advanced

analytics on routinely recorded data, we identified groups of herds with different AW status and good welfare was associated with higher production and longevity. :