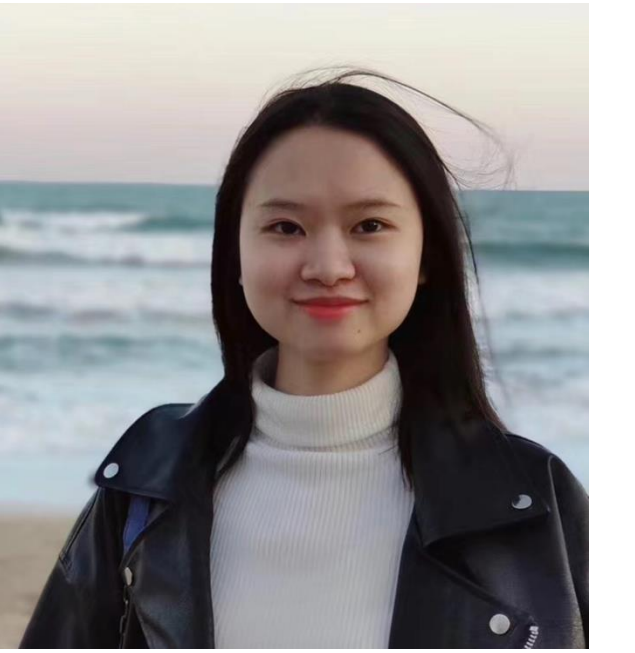
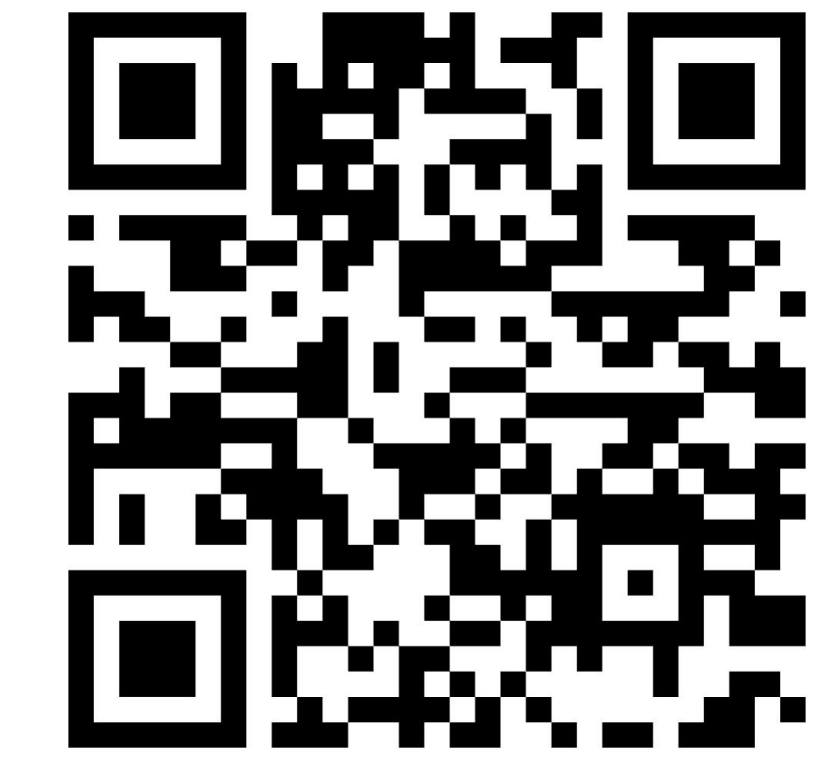


# Forecasting early-lactation diseases in Holstein dairy cows using milk spectra and machine learning

D. Lin<sup>1</sup>, J. Li<sup>1\*</sup>, J. A. Seminara<sup>2</sup>, J. A. A. McArt<sup>2\*</sup>

<sup>1</sup>Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong, China

<sup>2</sup>Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, NY, 14853, USA



Abstract #: 2621

## Introduction

Can we use raw milk spectra and algorithms to monitor health of early-lactation cows?



- Dairy cows commonly experience health disorders in the early lactation period.
- Blood sampling and analysis is costly for farmers and invasive for cows. Thus, it's necessary to develop new methods.
- Milk sampling and analysis is cost-effective and non-invasive, but its potential in forecasting early-lactation diseases has yet to be fully explored.

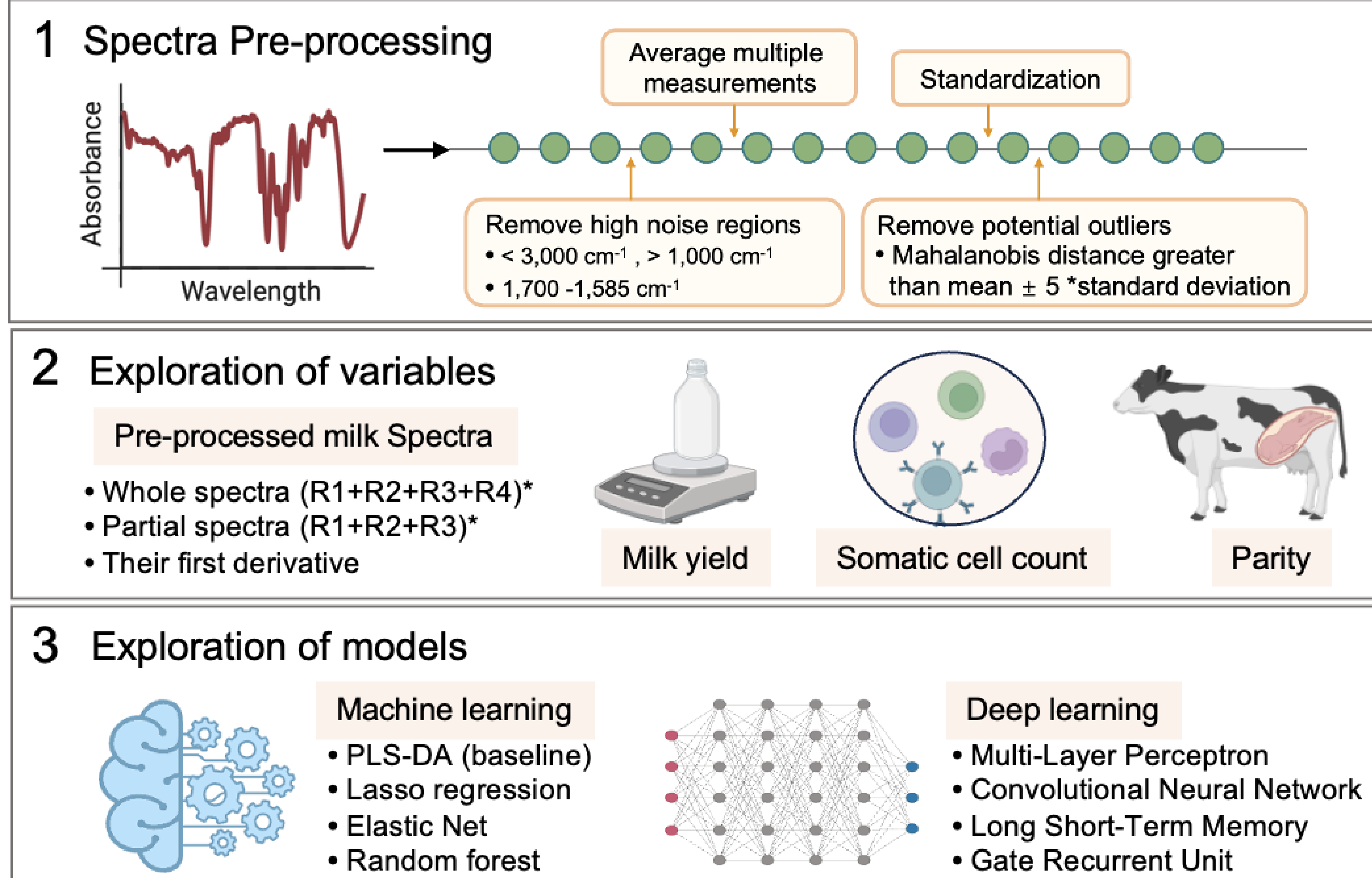
## Objectives

To fully explore the potential of milk FTIR spectra to forecast early-lactation disease in dairy cows, we set up two specific objectives:

- To evaluate whether milk spectra become more abnormal and more beneficial for prediction as disease diagnosis nears;
- To explore the performances of various machine learning and deep learning methods with some cow-level variables in forecasting disease diagnosis (metritis, displaced abomasum, ketosis, and mastitis) within the first 30 DIM.

## Materials & Methods

- Dataset: 1,144 cows and 6,113 milk samples on 1 commercial herd in New York from June through August 2021



## Conclusions & Implications

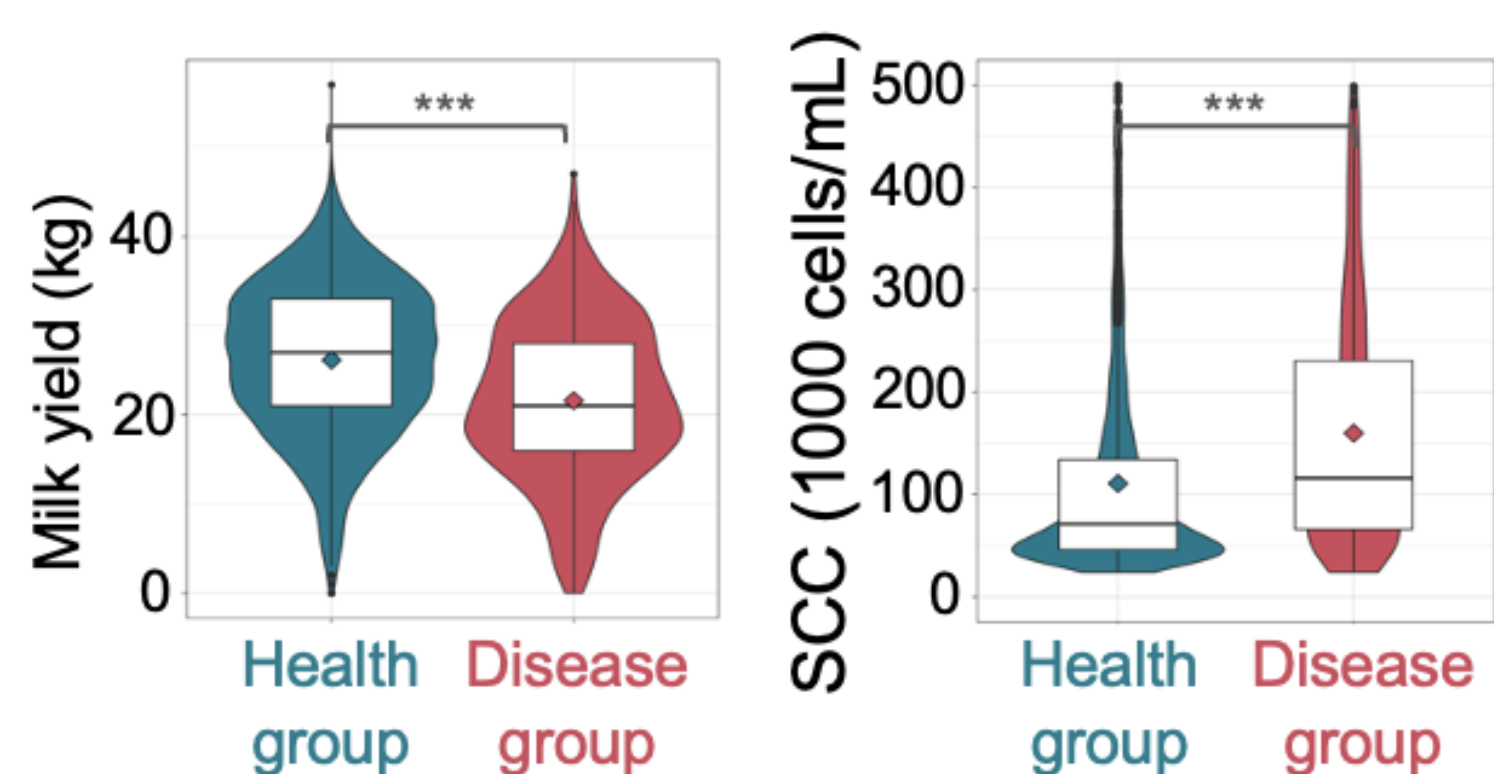
- Progressive changes in spectral regions related to the absorbance of fat, protein, and lactose are correlated with disease progression.
- Progressive changes in specific spectral regions enhanced predictive accuracy with disease progression.
- Predictive performance peaked in all models at 1 d prior to disease diagnosis.
- Including milk yield, somatic cell count, and parity into the milk FTIR spectra-based models improved the predictive accuracy.
- Machine and deep learning models outperformed the baseline PLS-DA.



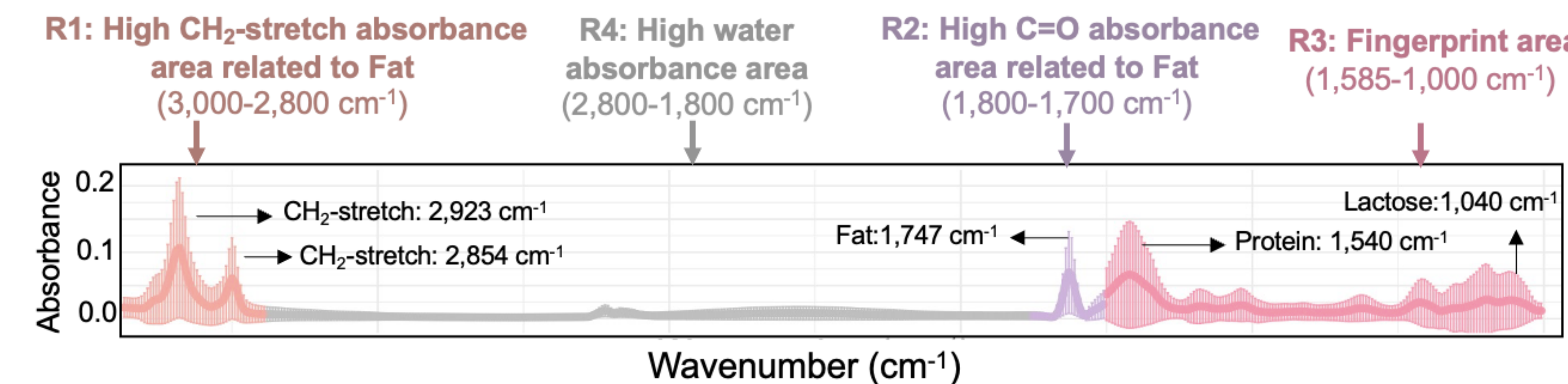
**This study highlights the effectiveness of integrating raw milk FTIR spectra with cow-level variables to forecast health conditions in early-lactation Holstein dairy cows via machine and deep learning models.**

## Results

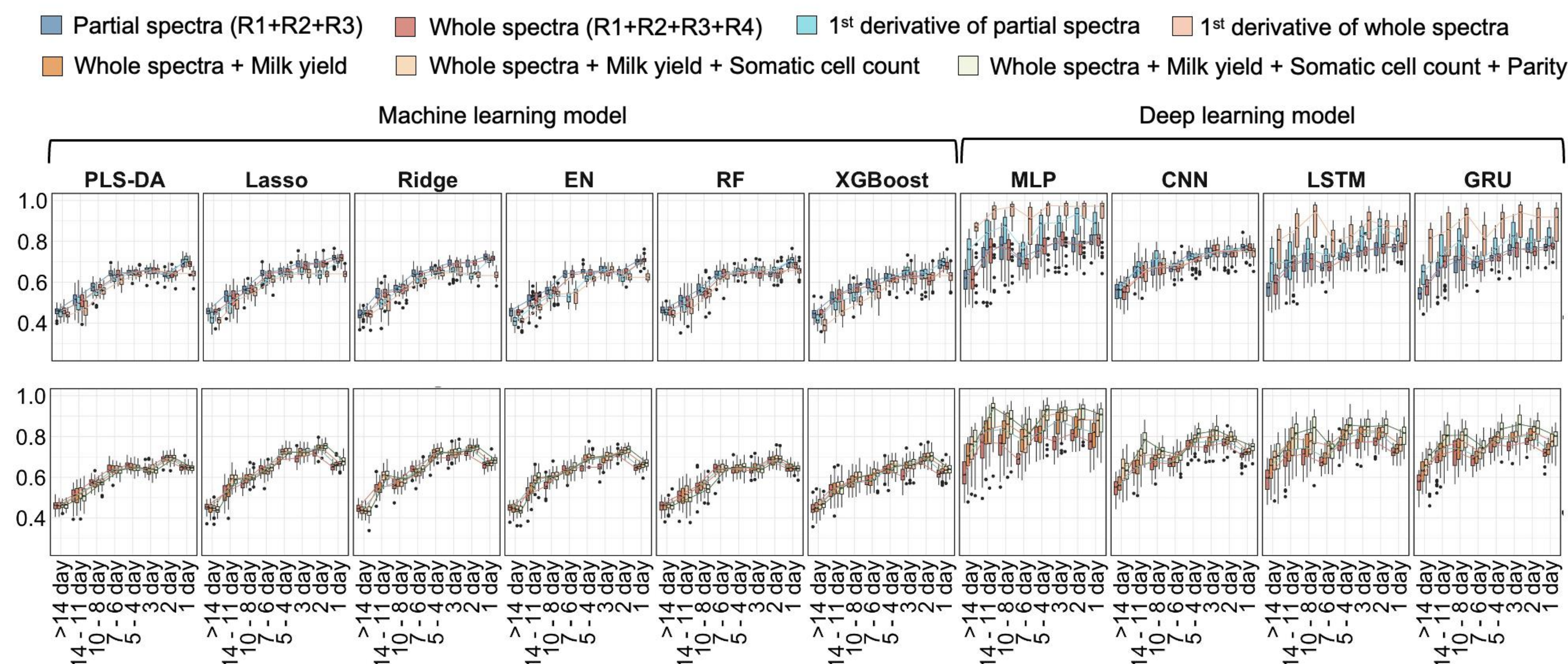
**Figure 1.** Descriptive statistics for cow-level variables in health group (n=825) and disease group (n=289).



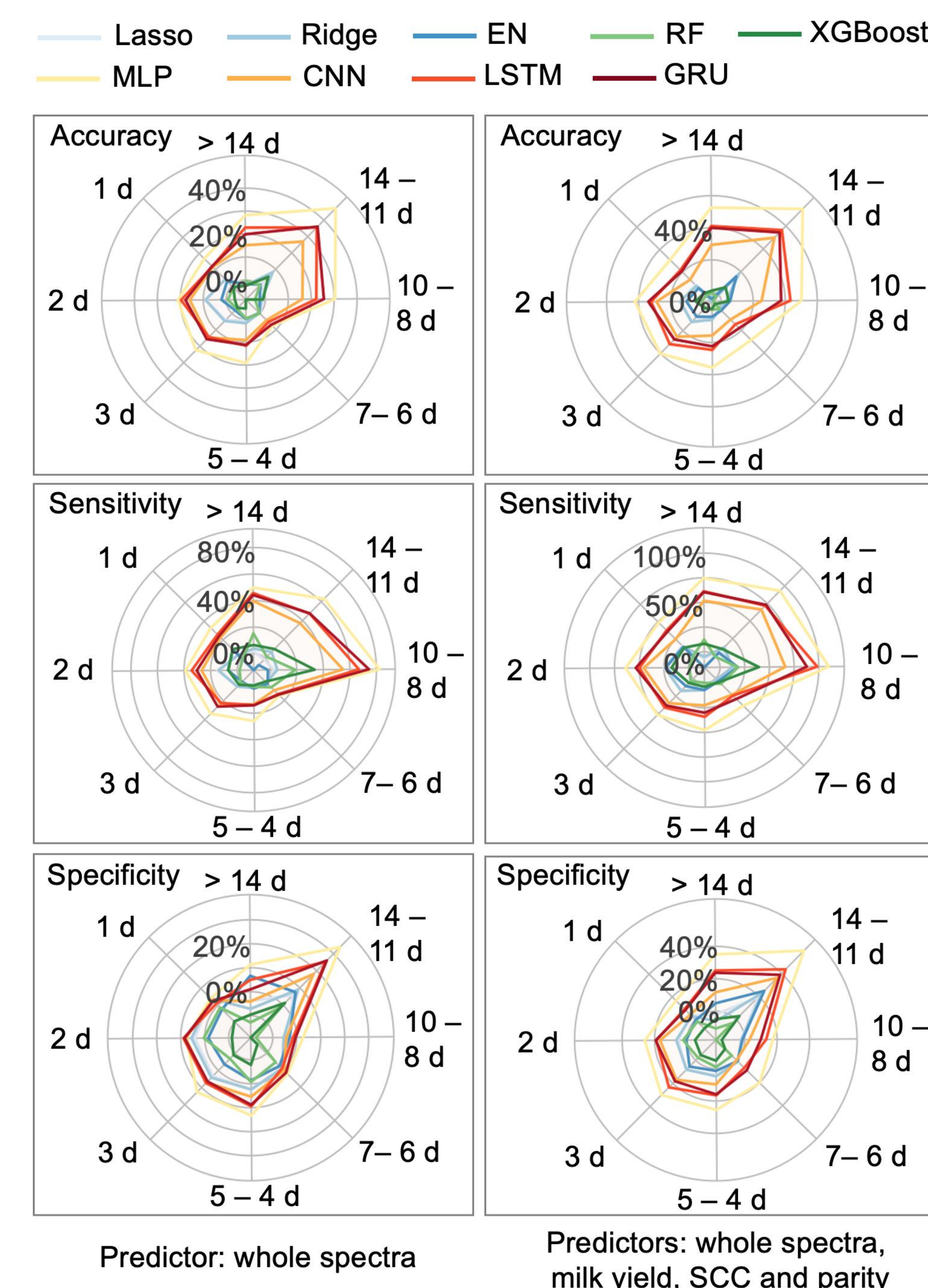
**Figure 2.** Absolute difference (mean  $\pm$  SD) in the pre-processed milk FTIR spectra (wavenumber ranging between  $3,000-1,000 \text{ cm}^{-1}$ , mean  $\pm$  SD) and four subregions (R1, R2, R3, R4) between health group and disease group.



**Figure 3.** Comparison of predictive accuracy between models using different variables on different times prior to disease diagnosis ( $>14 \text{ d}$ ,  $14-11 \text{ d}$ ,  $10-8 \text{ d}$ ,  $7-6 \text{ d}$ ,  $5-4 \text{ d}$ ,  $3 \text{ d}$ ,  $2 \text{ d}$ , and  $1 \text{ d}$ ), based on cross-validation.



**Figure 4.** Predictive performance compared to PLS-DA of models using whole spectra, milk yield, SCC, parity as the predictors on different times prior to disease diagnosis ( $>14 \text{ d}$ ,  $14-11 \text{ d}$ ,  $10-8 \text{ d}$ ,  $7-6 \text{ d}$ ,  $5-4 \text{ d}$ ,  $3 \text{ d}$ ,  $2 \text{ d}$ , and  $1 \text{ d}$ ), based on cross-validation.



**Figure 5.** Visualization of four representative cows. As the disease diagnosis day neared, there was a steady escalation in the models' predictive probability for a subsequent disease, reaching its peak on 1 d prior to disease diagnosis.

