Spectral sensors and a novel multiblock data fusion approach for barley pre-harvest germination discriminant analysis

Sebastian Helmut Orth1, Federico Marini1,2, Glen Patrick Fox 1,3, Stefan Hayward1, Marena Manley1

1 Department of Food Science, Stellenbosch University, Private Bag X1, Matieland, Stellenbosch, 7602, South Africa, sorth@sun.ac.za, stefanh@sun.ac.za, mman@sun.ac.za

2 Department of Chemistry, University of Rome “La Sapienza”, P.le Aldo Moro 5, 1-00185 Rome, Italy, federico.marini@uniroma1.it

3 Food Science and Technology, University of California Davis, 1Shields Ave, Davis, CA 95616, USA,

gpfox@ucdavis.edu

\*Corresponding author: Stefan Hayward (stefanh@sun.ac.za)

During every growing season, pre-harvest germination of malting barley (*Hordeum vulgare* L.), leads to major economic losses globally. Culminating toward a breaking point, where a new, unbiased and rapid detection method, capable of early stage germination is required to fill a gap in industry. Indeed, malting barley is cultivated for the purpose of controlled malting, where barley grain undergoes a series of steeping phases followed by controlled germination. Thus, barley damage through pre-harvest germination, a process triggered by untimely germination due to adverse environmental factors, is a major role player in detrimental and non-uniform malting. In addition to the detrimental nature, early-stage detection of pre-harvest germinated grain, using conventional rheological methods is not easy and mostly not possible, whilst also not including sample variation to justify grading decisions. The use of spectral imaging, to solve this specific industry problem, was investigated. Spectral information was obtained from two independent hyperspectral camera systems in the visible near-infrared and shortwave infrared waveband regions. Bringing about the necessity to maximise the spectral information available, a multiblock method was deemed viable. Thus, novel orthogonalised multiblock and multiblock variable selection classification methodologies, sequential and orthogonalised partial least squares linear discriminant analysis (SO-PLS-LDA) and sequential and orthogonalized covariance selection-linear discriminant analysis (SO-CovSel-LDA), were used. The aim, to ultimately classify between early stage, germinated and non-germinated single barley kernels, to capture intra- and inter sample variation. Classification accuracy of SO-PLS-LDA was 99.76% and when selecting and limiting the number of variables to 8, using SO-CovSel-LDA, an accuracy of 97% was obtained. This truly novel approach showed how VNIR and SWIR spectral information and a multiblock method is viable for industry and research applications. Further, orthogonalised data fusion allowed for modelling with a high degree of precision and sensitivity, without introducing redundant spectral variables commonly found in conventional low-and-mid-level approaches.

**Keywords:** Spectral imaging, SO-PLS-LDA, SO-CovSel-LDA, Single kernel analysis, Data fusion

**Acknowledgements:** The authors gratefully acknowledge receiving funding from ABInBev Pty Ltd.