Postharvest Biology and Technology 189 (2022) 111893

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

Segregation of 'Hayward' kiwifruit for storage potential using Vis-NIR spectroscopy

Mo Li [*], Reddy Pullanagari, Ian Yule, Andrew East

Massey AgriFood (MAF) Digital Lab, Massey University, Palmerston North, New Zealand

A R T I C L E I N F O

Keywords:
Actinidia deliciosa
Storability
Near infrared
Prediction
Machine learning

A B S T R A C T

Kiwifruit are often harvested unripe and kept in local coolstores for extended periods of time before being marketed. Many pre-harvest factors contribute to variation in fruit quality at harvest and during coolstorage, resulting in the difficulty in segregating fruit for their storage potential. The ability to forecast storage potential, both within and between populations of fruit, could enable segregation systems to be implemented at harvest to assist with inventory decision making and improve profitability. Visible-near infrared (Vis-NIR) spectroscopy is one of the most commonly used non-destructive techniques for estimation of internal quality of kiwifruit. Whilst many previous attempts focused on instantaneous quantification of quality attributes, the objective of this work was to investigate the use of Vis-NIR spectroscopy utilised at harvest to qualitatively forecast storage potential of individual or batches of kiwifruit. Commercially sourced 'Hayward' kiwifruit capturing large variability of storability were measured non-destructively at harvest using Vis-NIR spectrometer, and then assessed at 75, 100, 125 and 150 days after coolstorage at 0 °C. Machine learning classification models were developed using at-harvest Vis-NIR spectral data, to segregate storability of kiwifruit into two groups based on the export FF criterion of 9.8 N. The best prediction was obtained for fruit stored at 0 °C for 125 days: approximately 54% of the soft fruit (short storability) and 79% of the good fruit (long storability) could be predicted. Further novelty of this work lies within an independent external validation using data collected from a new season. Kiwifruit were repacked at harvest based on their potential storability predicted by the developed model, with the actual post-storage performance of the same fruit assessed to evaluate model robustness. Segregation between grower lines at harvest achieved 30% reduction in soft fruit after storage. Should the model be applied in the industry to enable sequential marketing, significant costs could be saved because of reduced fruit loss, repacking and condition checking costs.

1. Introduction

The kiwifruit industry of New Zealand has an export value of NZ$2.3 billion, making up a third of the total value of NZ horticultural sector (Aitken and Warrington, 2020). In commercial practices, kiwifruit are often harvested unripe and kept in coolstorage for months before being exported to distant markets. During postharvest storage, kiwifruit ripen, and this results in physiological and biochemical changes such as starch-to-sugar conversion, and the softening of the flesh (Beever and Hopkirk, 1990). Flesh firmness (FF) is one of the most important postharvest quality indices for kiwifruit, determining the storage potential and eating quality of kiwifruit. As part of quality control measures, fruit are tested based on the minimum export FF criterion (1 kg_f or 9.8 N) using destructive methods prior to shipping, and again on arrival at the distant market (East et al., 2013). However, such destructive testing can only be applied to a small portion of the fruit in a discrete sampling plan because it leads to direct fruit loss. Removal of over-soft fruit that are unsuitable for sale incurs significant costs.

The variability in fruit at the point of harvest contributes to a wide range of storage potential. A kiwifruit grower line (GL) represents a batch of fruit that are sourced from a set of unique growing conditions (i.e. orchard location, preharvest treatments, orchard, harvest and postharvest operations, maturity at harvest etc.). Grower line is one of the main factors influencing variability (Buxton, 2005; Patterson, Currie, 2011). Although a large percent of the fruit would store well through the season, it is the poorest-storing fruit in a line that influences the storability of the line (Tanner et al., 2012). It would be beneficial to utilise the variability in the population and segregate fruit with different

* Corresponding author.
E-mail address: m.li2@massey.ac.nz (M. Li).

https://doi.org/10.1016/j.postharvbio.2022.111893
Received 12 October 2021; Received in revised form 25 February 2022; Accepted 25 February 2022
Available online 28 March 2022
0925-5214/

intended storability in the supply chain, in order to improve the efficiency of packing operations and reduce postharvest cost through fruit loss. There are two potential segregation systems (Fig. 1): within and between grower line (GL). Within GL segregation (Fig. 1a) identifies individual short-storing fruit and aims to separate these from the long-storing ones within the same batch. The outcome of this segregation would be two groups of fruit with the poor-storing fruit from each GL being separated and shipped earlier in the season. For between GL segregation (Fig. 1b), poor-storing lines are identified as they contain a larger number of potentially short-storing fruit and hence have, on average, lower storability. Segregation would result in these GLs being separated from the population for earlier sale. Both systems would be useful to assist with inventory decisions for sequential marketing but would require different implementation and have different outcomes on postharvest performance and grower orchard gate return.

Identifying long-storing fruit is important for making inventory decisions because firmness decreases rapidly for all the fruit from ~80–15 N during the first 70 days of storage (Beever and Hopkirk, 1990). During this period the proportion of soft fruit is very low and majority of the fruit would have already been shipped. Afterwards fruit continue to soften slowly during 100 – 175 days of storage and the average firmness will be approaching and eventually go below 9.8 N or 1 kgf (Beever and Hopkirk, 1990; Jabbar, 2014), causing soft fruit to become prominent and problematic. As a result, the higher proportion of soft fruit would accelerate further fruit softening in the remaining batch. The ability to predict and segregate fruit with storability beyond 100 days would not only enable the reduction of direct fruit loss resulted from short-storing fruit becoming unacceptable for export, but also enable the separation of long-storing fruit from an ethylene environment produced by short-storing fruit during softening (Samarakoon, 2013), preventing secondary fruit loss (Jabbar and East, 2016).

The use of visible-near infrared (Vis-NIR) spectroscopy to non-destructively assess internal quality of fresh produce has been well established in postharvest applications (Nicolaï et al., 2007a; Walsh et al., 2020a). The Vis range of the spectrum (400 – 750 nm) provides information on colour and pigment analysis (e.g. chlorophylls, carotenes) of the target object, whereas the NIR range (750 – 2500 nm) is associated with the macro constituents (e.g. water and carbohydrate) of the product (Walsh et al., 2020a). The attenuation spectrum is the result of both absorption and scattering of photons from a 'spot' measurement, and usually possesses high multicollinearity because the signal at a specific wavelength can be highly correlated with those at neighbouring wavelengths. Therefore, multivariate statistical techniques (also known as chemometrics) are required to reduce this multicollinearity and extract the information relevant to quality attributes which is hidden in the spectrum (Nicolaï et al., 2007a; Saeys et al., 2019). Both linear and non-linear methods have been used for quantification (regression) and

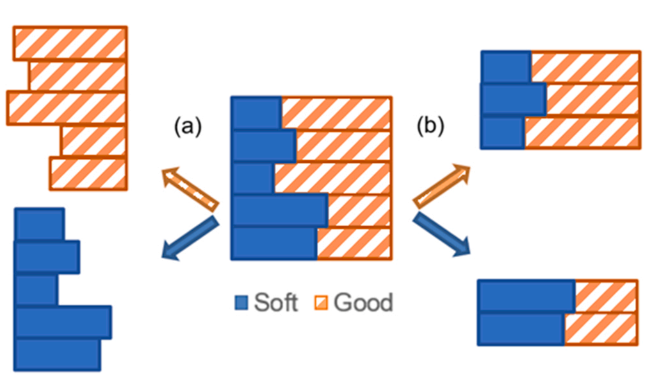

Fig. 1. Conceptual diagram of segregation of five batches of kiwifruit (a) within grower line and (b) between grower line. Hatched orange colour indicates good-storing fruit/lines whereas solid blue colour indicates poor-storing fruit/lines.

qualification (classification) of quality attributes. Partial least squares regression (PLSR) is the most common multivariate linear regression method (Walsh et al., 2020a), whereas recent advancement in kernel functions and neural networks provide a means to deal with non-linearities in the data. For classification problems, linear discriminant analysis (LDA), sometimes used with PLS (PLS-DA) is a common method, although more recently, machine learning techniques such as support vector machines (SVM) have also been used and can achieve robust calibration models with good prediction outcomes.

For any spectroscopy technique, calibration is a process which develops the mathematical relationship, in the form of a model, between measured sample properties and the attenuation spectra of the set of known reference samples (Zeaiter et al., 2005). For biological samples such as fresh fruit, the model usually requires data from multiple seasons and/or origins to include a wide range of inherent variability, in order to ensure robustness of the model. Once an NIR instrument has been calibrated against a reference method for the measurement of a particular sample property, it can be used to predict unknown samples and the prediction errors can be estimated, a process known as validation. Validation of calibration provides the basis for calculation of true measurement error by comparing NIR measurements to reference method measurements on a new set of samples (Williams and Norris, 1987). If the two are essentially the same, the model provides accurate prediction and will be useful for future predictions (Sun, 2009). The validation process can be internal and external depending on the nature of the validation dataset. Internal validation utilises dependent datasets that derive from the calibration set through techniques such as leave-one-out or the holdout method, whereas external validation typically uses a test set that is independent from the calibration dataset (ideally from different seasons and/or origins) and provides more realistic assessment of the model robustness. According to Walsh et al. (2020a), approx. 90% of the publications on the use of NIR in postharvest applications either only adapted internal cross-validation methods, or deliberately selected test sets which highly resemble the calibration set, resulting in arguably over-optimistic model performance compared to the use of an independent test set which is representative of implementation in a practical scenario.

For the kiwifruit industry, Vis-NIR spectroscopy is commonly used in quantitative analysis of quality attributes by developing regression models. Focus has been mainly on instantaneous estimation of at-harvest quality attributes. The success in estimation of sugar content of the fruit [e.g. total soluble solids (TSS), and dry matter content (DMC)] has been demonstrated in the literature (McGlone and Kawano, 1998; Osborne et al., 1999; Schaare and Fraser, 2000; Clark et al., 2004; Moghimi et al., 2010), and has also eventuated in the adaptation of multilane NIR sensors fitted to a commercial kiwifruit grader (Inspectra^2, TOMRA Systems ASA, Bay of Plenty, NZ) for sorting of kiwifruit according to DMC. On the other hand, estimation of flesh firmness has been more challenging (McGlone and Kawano, 1998; Costa et al., 1999; Lee et al., 2012), primarily due to the minimum chemical compositional changes taking place during firmness progression in kiwifruit, and the large variance in the physical measurements of firmness which reportedly contributed to up to 80% of the observed RMSE in regression models (Li et al., 2017). However, cellular structural changes such as cell wall degradation could affect scattering of light which may still be detectable by the NIR signal, and firmness can be related to changes of other attributes such as water content, providing opportunities for secondary correlations (Walsh et al., 2020a).

From a practical point of view, sorting of samples often requires a qualitative analysis (i.e. classification). A few previous attempts have been conducted to evaluate the ability of NIR to perform discriminant prediction using classification models. The closest example is that of Feng (2003) where NIR spectral and fruit quality data were collected at harvest to classify 'Hayward' kiwifruit for storage potential using canonical discriminant analysis (CDA). At-harvest NIR spectra were calibrated with various at-harvest fruit attributes, and the calibrated model was then used for the prediction of post-storage firmness, allowing for segregation of disordered fruit from healthy ones. The classification accuracy was 67% and 35% for healthy fruit and fruit with soft patch respectively. A similar approach was described in Feng et al. (2014) who attempted to segregate storage potential of individual kiwifruit based on at-harvest NIR spectra but the model required *a priori* knowledge of at-harvest attributes (e.g. skin and flesh colour, FF, TSS and DMC) and in-storage acoustic firmness measurements. For a targeted 75% true positive rate (rejected fruit accurately classified), the false positive rate (good fruit classified wrongly) was 30% and 40% respectively for 'Hayward' and 'SunGold™' when logit-boost decision stumps were used for classification model development. The moderate false positive rates were considered acceptable for industry application because fruit classified as reject at harvest could still be marketed earlier in the season, rather than being disposed of, to prevent fruit loss. Previously, Clark et al. (2004) had categorised chilling injury (CI) propensity of 'Hort16A' kiwifruit based on at-harvest Vis-NIR reflectance using CDA classification, and obtained 52–66% accuracy for disordered fruit (with rots and CI after storage), and 80–89% accuracy for good fruit. This would indicate a reduction in disorder incidence from 33.9% to 17.9% and 14.7–8.5% for both harvests. However, external validation was not conducted. Burdon et al. (2014) used at-harvest NIR spectral data calibrated with at-harvest attributes to predict the incidence of CI of stored 'Hort16A' kiwifruit, and concluded that the classification was not useful due to a large orchard factor which contributed to considerable variation in the minimum or maximum threshold for the development of CI.

In other crops, prediction of the storability of apricots was carried out by fitting an exponential model to describe the relationship between FF_0, the at-harvest FF estimated by Vis-NIR spectral data, and FF_{Final}, post-storage firmness predicted based on FF_0 (Feng et al., 2013). Segregation of storage potential was based on the projected FF_{Final} in comparison to the minimum standard for sale at retail (10 N). This segregation provided theoretical limits for initial firmness of apricot for the two cultivars studied but validation of the segregation model was not conducted. In addition, Zude et al. (2006) discriminated post-storage quality levels of apple as a result of storage condition treatments using non-destructively estimated TSS (Vis-NIR) and FF (acoustic impulse resonance frequency) values measured after storage, and obtained 77% and 84% overall accuracy respectively.

To date most relevant studies required calibration of the model with at-harvest attributes in addition to spectral data in order to perform prediction or segregation of future storability. However, such an approach requires destructive sampling of the fruit, and the accuracy for prediction is inevitably affected by sample selection during this calibration process. In some cases in-storage quality measurements were required (e.g. Feng et al., 2014) which would require extra labour reshuffling pallets of fruit in order to obtain a representative sample. In addition, most relevant studies reported model performance using internal validation and the robustness of the model to perform prediction on an independent data set was not determined. In this work, an alternative approach has been proposed to perform prediction/classification prior to storage directly using NIR spectra data as the only input. In this case a black-box model will be developed using supervised learning. In kiwifruit, the final firmness is influenced by a range of fruit characteristics at harvest which affect both maturity stage and rate of softening at harvest. It is hypothesised that the spectra capture this initial characteristic of the fruit which may be indicative of the potential storage outcome of the same fruit. The resulting model would need to be validated externally to demonstrate robustness beyond the current calibration population.

Therefore, this work, to the best of the authors' knowledge, is the first attempt to utilise Vis-NIR spectral data as the sole input at harvest, to perform qualitative prediction of kiwifruit firmness retention properties after coolstorage by developing binary classification models. Machine learning algorithms will be selected to develop a suitable model which aims to segregate individual kiwifruit into two groups based on the export firmness criterion (FF $\geq$ 9.8 N). The developed model will be applied to an independent test set as an external validation.

2. Materials and methods

2.1. Calibration model development

2.1.1. Calibration data set

To investigate the feasibility of Vis-NIR spectroscopy for prediction of storage potential of 'Hayward' kiwifruit, four sets of Vis-NIR spectral and fruit quality data were collected as described in Section 2.1 of Li et al. (2017) and used for calibration. The kiwifruit GLs were sourced from multiple regions over two seasons (2012 – 2013) capturing the wide variability of fruit storability observed within the industry (Table 1), which is the prerequisite for building a robust model. For all the datasets, the at-harvest Vis-NIR spectral data were collected non-destructively, and the post-storage FF of the same fruit were measured at different times (75 – 150 days after storage) to investigate the model performance for those storage periods (Fig. 2).

2.1.2. Spectral measurements

A commercial full-range Vis-NIR spectroscopy system (FieldSpec® Pro, Malvern Panalytical., Malvern, UK) was used. The instrument used three types of detectors to cover both the visible and the NIR range of the spectrum, a silicon detector (350–1000 nm) and two InGaAs detectors measuring the shortwave infrared range (1000–1800 nm and 1800–2500 nm respectively). The optical fibre of the instrument was coupled with a probe (Hi-Brite, Malvern Panalytical., Malvern, UK) for contact measurements with a spot size of 10 mm in diameter. The probe was fitted with a high intensity halogen lamp to produce consistent illumination in a broad electromagnetic spectrum. The probe was fixed to a burette stand in a nadir position and connected to the instrument through an optical fibre cable. A diffuse reflectance material (Spectralon®, Labsphere Inc., North Sutton, USA) panel was used as a reflectance standard and to convert raw spectra to reflectance. At the time of scanning, each fruit was measured at two equatorial locations (90° apart) corresponding to the locations of post-storage flesh firmness measurements. Spectra were recorded in 1 nm internals with spectral resolution of 3 nm (at 700 nm), 10 nm (at 1400 nm) and 12 nm (at 2100 nm).

2.1.3. Fruit quality measurements

Fruit flesh firmness (N) was measured using an electronic QALink Penetrometer (Willowbank Electronics Ltd., Napier, New Zealand) fitted with the standard 7.9 mm Magness-Taylor probe. Measurements of peak penetrating force were made at two equatorial locations (90° apart) after removal of a thin layer (1 mm) of the fruit skin. The penetration speed was 20 mm s^{-1} and the puncture depth was 8 mm.

2.1.4. Data analysis

2.1.4.1. Spectral pre-treatments. The raw spectral data were pre-

Table 1
Summary of Vis-NIR spectral and post-storage firmness data sets collected in 2012 – 2013 for developing and validating classification models. Numerical values represent the number of fruit measured at each time of storage. Soft and Good means flesh firmness of the fruit is < 9.8 N and ≥ 9.8 N respectively.

Storage time (day)	Calibration					
	Season				Total	
	2012	2012	2013	2013	Soft	Good
75	255	40	40	0	**35**	**300**
100	255	40	240	320	**87**	**768**
125	255	40	40	0	**121**	**214**
150	0	40	240	320	**133**	**467**

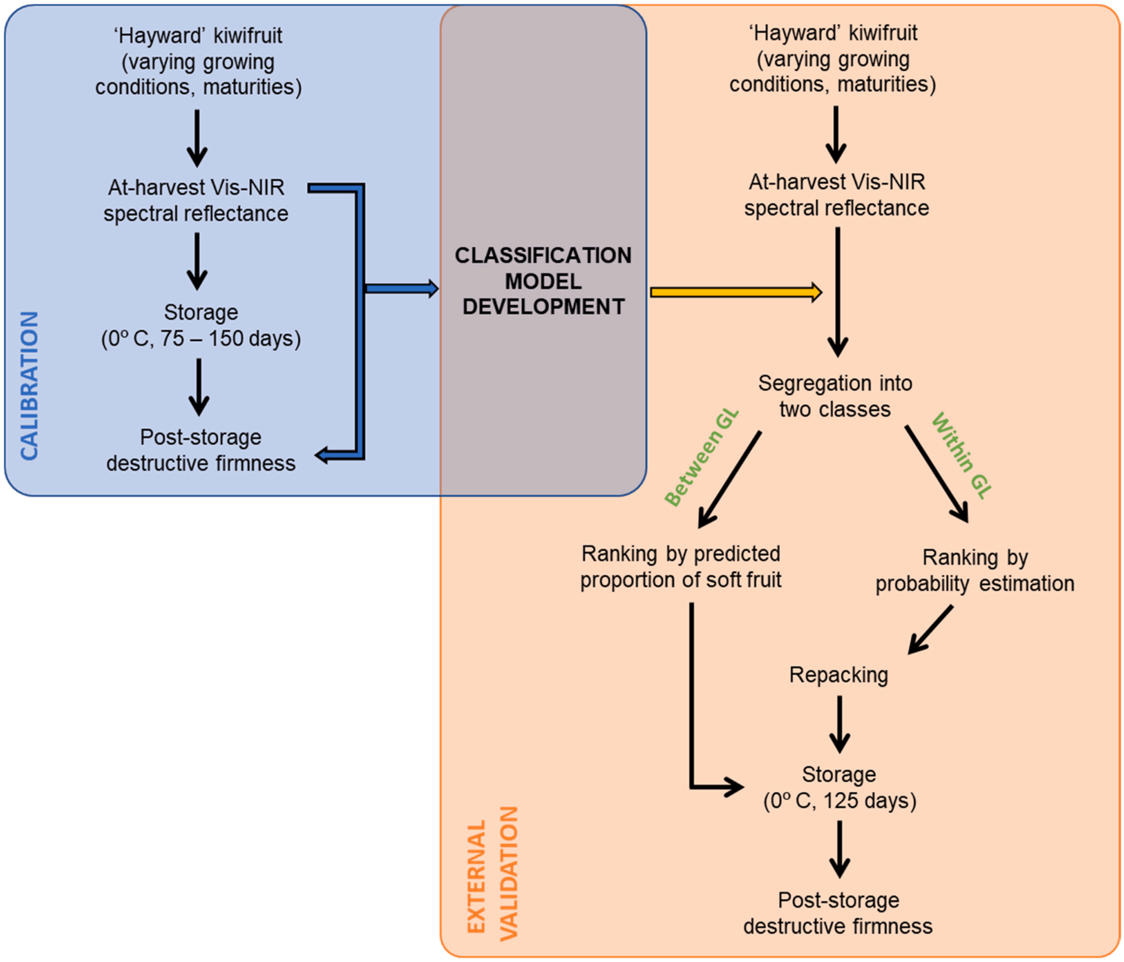

Fig. 2. A schematic illustrating the process of classification model development and external validation processes carried out in the current study. Calibration was carried out using data collected in 2012–2013 whereas external validation was carried out using data collected in 2015. GL denotes grower lines.

processed using The Unscrambler® (Version X10.3; CAMO Software AS., Oslo, Norway). Spectral data from all experiments were first truncated to 400–2450 nm to eliminate high noise regions at both ends. Reflectance was converted to absorbance via Log transformation. First order derivation using a Savitzky-Golay smoothing algorithm (Savitzky and Golay, 1964) was then applied to reveal the hidden information in the spectra as well as to reduce the noise in the data. Spectra were then normalised sample-wise using a unit vector with a length of 1, so that all the data were the same scale.

2.1.4.2. Model calibration and internal validation. Fruit were categorised into two storage potential groups based on their FF values after cool-storage: Soft (FF $<$ 9.8 N; too soft to export) and Good (FF $\geq$ 9.8 N; good to dispatch). The predictive relationship between at-harvest spectral data and post-storage fruit grouping was investigated by developing a classification model for each of the four post-storage time points. A ten-fold cross-validation method was used for internal validation.

Machine learning algorithms were employed to identify the most suitable method for classification model development. Data processing was performed using the Weka software package (Version 3.7.12; University of Waikato, Hamilton, New Zealand; Hall et al., 2009).

2.1.4.2.1. Random forests. A decision tree finds features in the input variables and identifies the threshold for the features that best splits the data into separate classes (Quinlan, 1986). Each feature attribute is presented as a node in the tree, with each possible threshold of each attribute as a branch and a class label as each leaf. However, this method is prone to over-fitting and has high variance. Random Forests are an ensemble of decision trees. In this method each model (tree) is trained independently using a random small subset of features for the split. As a result the predictions from the sub-trees are uncorrelated or weakly correlated, resulting in lower variance (Nguyen et al., 2006). The generalisation error for forests converges to a limit as the number of trees in the forest becomes large. Prediction is made by aggregating majority vote for the predictions of the ensemble (Anne-Michelle and Mousumi, 2007). The algorithm of RF was developed by Breiman (2001):

$$\{T_1(X), \ldots, \quad T_J(X)\} \tag{1}$$

where $X = \{x_1, \ldots, x_p\}$ is a p-dimensional vector of a sample or properties associated with a sample. The ensemble produces J outputs $\{\widehat{Y}_1 = T_1(X), \ldots, \widehat{Y}_J = T_j(X)\}$ where $\widehat{Y}_j$, $j = 1, \ldots, J$, is the prediction for a sample by the jth tree. Outputs of all trees are aggregated to produce one final prediction, $\widehat{Y}$. For classification problems, $\widehat{Y}$ is the class predicted by the majority of the trees (Svetnik et al., 2003).

2.1.4.2.2. Support vector machine classification. Support vector machines (SVM) with sequential minimal optimisation (SMO; Platt, 1999) was used for classifying the spectral data. For any set of two-class objects, the SVM finds the unique hyperplane having the maximum margin for optimal discrimination. The hyperplane defines the borders for each class with specific objects within the class, and these objects are referred to as support vectors (Ivanciuc, 2007). The support vectors are used to classify the samples. For non-linear classification, the coordinates of the input objects are mapped into a high-dimensional feature space using different kernel functions. The kernels can be computed in the same space as the input objects allowing linear algorithms to work with higher dimensional feature space. Classification is accomplished by a weighted sum of kernels evaluated by the support vectors (Ivanciuc, 2007). The SMO algorithm was used to speed up the training of SVMs by reducing a large quadratic programming optimisation problem into a series of small optimisations (Mohri et al., 2012). Suppose there are N data points in the training dataset,

$$\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \tag{2}$$

where $x_i \in \mathbb{R}_N$and $y_i \in (+1, -1)$.

Consider a hyperplane defined by (w, b), where w is a weight vector and b is a bias. A new object x can be classified with:

$$f(x) = sign \quad (wx + b) = sign \quad (\sum_i^N a_i y_i (x_i x_j) + b) \tag{3}$$

where (x_i, x_j) is a set of training data points and a_i is the Lagrange multipliers which is minimised with respect to w and b and maximised with respect to $a_i \geq 0$ (Gunn, 1998; Pachghare and Kulkarni, 2011).

For real-world data, the common approach is to solve the classification using a soft margin, meaning that the hyperplane separates most but not all of the data points. In this case the soft slack variable, ξ_i and the capacity constant, C will be required:

$$\text{minimise}\frac{1}{2}\|w\|^2 + C\sum_i \xi_i \tag{4}$$

$$\text{subject to}\left\{ \begin{array}{c} y_i f(x_i) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{array} \right\} \tag{5}$$

In this study the RBF kernel was used. This function generates hidden units that represent the coordinates of the objects in the input space (NIR spectral data and FF values). The output of an object for a given instance (the class that the sample belongs to) depends on the distance between the object and its instance. This distance is converted into a non-linear measure. The hidden units are referred to as RBFs and the hyperplane is formed when a given hidden unit for the objects in the instance space produces the same outputs (Witten et al., 2011). The algorithm for RBF kernels is:

$$\text{Kernel}(x_i, \quad x_j) = exp(-\gamma \| x_i - x_j \|^2) \tag{6}$$

where γ is the variable parameter (Gunn, 1998).

The robustness of the calibration model could be optimised by finding the optimal constant C and kernel parameter γ. However due to the limitation of the software Weka and time constraints, a grid search of the least RMSEs was not conducted. The default values (C = 1 and γ = 0.01) were used for developing all the models.

2.1.4.2.3. Boosted decision stumps. Decision stumps (DS) are one level decision trees with two terminal nodes (Friedman et al., 2000). In this method, each node in a DS represents a feature in a sample, and each branch represents a threshold value that the node can take. Samples are classified starting at the root node and sorting them based on their feature values (Kotsiantis et al., 2006). Stumps are weak leaners and usually have low variance but high bias (Friedman et al., 2000).

Boosting algorithms were first introduced by Freund and Schapire (1996) to provide a way of combining performance of many weak classifiers to produce a powerful committee. It uses a sequential algorithm in which each new weak learner is built based on the performance of the previously generated predictors (Jung, 2009).

AdaBoost algorithm (Freund and Schapire, 1996) assigns equal weight to all samples in the training data. When a classifier is formed by the learning algorithm, the algorithm reweights each sample according to the prediction output. The weight of correctly classified samples is decreased and that of misclassified samples is increased. A new classifier is then built for the reweighted data and focuses on predicting the previously misclassified samples correctly. Once again the algorithm reweights the samples according to the new classifier. The weights after iteration reflect how often the samples have been misclassified. Whenever error on the weighted training data exceeds or equals 0.5, or equals 0, the boosting procedure deletes the current classifier and does not perform any more iteration. To predict a new sample, the output of a series of classifiers generated by the boosting method is combined using a weighted vote, where: (Witten et al., 2011)

$$weight = \quad - \quad \log\frac{e}{1-e} \tag{7}$$

where e denotes the classifier's error on the weighted data and is a fraction between 0 and 1 (Witten et al., 2011).

LogitBoost was first introduced by Friedman et al. (2000) for fitting additive logistic regression models by maximum probability. It computes 'response variable' that encodes the error of the currently fit model on the training examples in terms of probability estimates. LogitBoost decision stumps use the logit transform to translate the probability estimation problem into a regression problem, and solve the regression task using DS (Witten et al., 2011). The probability of a sample being class A is a number between 0 and 1. If the number is more than 0.5 the algorithm will categorise the sample as class A, and vice versa. The probability for each instance can be calculated as:

$$\text{p}\left(\frac{1}{\mathbf{a}}\right) = \quad \frac{1}{1 + \quad e^{-\Sigma \quad f_{j(\mathbf{a})}}} \tag{8}$$

where f_j is the jth regression model and $f_{j(\mathbf{a})}$ is its prediction for sample **a** (Witten et al., 2011).

2.1.4.3. Classification model performance evaluation. Model performance was evaluated amongst different storage time points and between the various classification algorithms chosen. This section describes the performance metrics selected for this evaluation.

For each of the storage time points, confusion matrix was produced for each storage potential group. In this matrix, true positive (TP) is referred to as correctly classified soft fruit (< 9.8 N). True negative (TN) is the correctly classified good fruit (≥ 9.8 N). False positive (FP) is the number of classified soft fruit which are good. False negative (FN) is the number of classified good fruit which are soft. For this study, the proportion of actual soft fruit in the segregated soft and good groups respectively, i.e. the TP and FN rates, were used to assess model robustness. This is because the TP rate represents the true correct classification of soft fruit and the FN rate indicates the potential fruit loss in the segregated good fruit population which is very important for justification of industrial applications.

For comparison amongst the various classification algorithms, additional parameters including classification accuracy, overall accuracy, kappa values, mean absolute error, recall and precision were estimated.

Classification accuracy represents the fruit with accurately predicted FF groups (based on the minimum export criterion of 9.8 N) as a percent of the actual fruit belonging to that FF group at a specific time point. Overall accuracy (OA) is the percentage of correct predictions in the entire population, i.e. (TP + TN) / n, where n is the total number of samples.

Kappa is a value that ranges between 0 and 1 which indicates the reliability of a classifier on a specific dataset. The closer the value is to 1, the more reliable the classifying algorithm is. The kappa statistic can be calculated using the equation:

$$\text{Kappa} = \frac{\text{Observed accuracy} \quad - \quad \text{expected accuracy}}{1 - \quad \text{expected accuracy}} \tag{9}$$

where observed accuracy is the total number of instances that were classified correctly throughout the entire confusion matrix; expected accuracy is defined as the accuracy that any random classifier would be expected to achieve based on the confusion matrix and can be calculated as:

$$\text{Expected accuracy} = \frac{\frac{(TP+FN)\times (TP+FP)}{n} + \frac{(TN+FP) \times (TN+FN)}{n}}{n} \tag{10}$$

where n represent the number of total observations.

Mean absolute error (MAE) is the mean of overall error made by the classifier.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\widehat{\theta}_i - \theta_i\right| \tag{11}$$

where $\widehat{\theta}_i$ is the predicted value and θ_i is the observed value.

Recall is the proportion of samples belonging to the positive class (i.e. Soft) that are correctly predicted, i.e. Recall = TP / (TP + FN).

Precision is the proportion of actual positive samples in the predicted positive class, i.e. Precision = TP / (TP + FP).

In addition, the performance of the classifiers was evaluated by obtaining the receiver operating characteristic (ROC) curves. The ROC curve is used to characterise the trade-off between hit rate (signal) and false-alarm rate (noise) over a noisy channel (Gorunescu, 2011; Witten et al., 2011; Walsh et al., 2020b). As a result the ROC curves can be used to select classifiers, based on their performance (Gorunescu, 2011). Often the ROC curves are plotted using the TP rates ('benefits') against the FP rates ('costs'). However for this study, the FN rates are more important as they are the true costs of a poor segregation (number of soft fruit in the predicted good batch). Therefore the ROC curves were obtained by plotting the TN values against the FN values. The samples were sorted in descending probability order according to the predicted probability of a true response. In ideal situations the curves should be as close to the upper left corner (vertical axis) as possible (Witten et al., 2011) because high TN rate results in consecutive vertical lines which will bring the curve to coordinates with low FN rate. The point (0, 1) represents perfect classification (i.e., no FNs), whereas a completely random guess would form a diagonal line (no discrimination) from the left bottom to the upper right corner (Gorunescu, 2011).

A simple way of evaluating ROC curves is to estimate the area under curve (AUC). A higher AUC value suggests better classification performance. Hence the AUC values were used to represent ROC performance for algorithm comparison. In general an AUC value between 0.8 and 1.0 indicates good to excellent classification accuracy, whereas 0.7 – 0.8 is considered fair accuracy. However, cautions should be taken when using AUC to evaluate model performances, because over-simplifying ROC curves into a single AUC number may lose information about the pattern of trade-offs of a particular classifier (Gorunescu, 2011).

Final ranking of classifiers was carried out using the Garrett's Ranking Technique (Garrett, 2002). It was calculated as percentage score using the equation:

$$\text{Percentage score} = \frac{100\quad (R_{ij} - 0.5)}{N_j} \tag{12}$$

where R_{ij} is the rank given for the i^{th} item j^{th} individual; N_j is the number of items ranked by j^{th} individual.

2.2. Segregating storage potential of kiwifruit using developed classification model – an external validation

Kiwifruit softening is known to be highly sensitive to ethylene, even in coolstorage conditions (Jabbar and East, 2016), with ethylene production of kiwifruit dramatically increasing as fruit soften (Samarakoon, 2013). Consequently, as short-storing fruit become soft, they have the potential to produce within pack an ethylene environment that softens otherwise long-storing fruit, reducing overall firmness in the same tray. This 'cross-contamination' effect can be greatly reduced if short-storing fruit can be identified and separated from long-storing fruit prior to storage. This is because, stored separately, long-storing fruit would go through normal softening in a relatively uncontaminated cooling environment with minimal interferences from soft fruit within the tray.

In external validation, physical separation of fruit was performed to facilitate the potential benefits of the pre-softening segregation. At-harvest Vis-NIR data was captured and analysed using the model developed in Section 2.1, and then individual fruit ranked on prediction of storability and resorted into trays based on this ranking, within each grower line. This process resulted in trays of fruit which are sorted by their storage potential as predicted from the Vis-NIR data, which were then stored in a coolroom, with FF measured after long term storage. At the same time the Vis-NIR data and resulting predictions from the data enable a prediction of the between grower line storability (Fig. 2).

2.2.1. Pre-storage Vis-NIR measurements

'Hayward' kiwifruit from a total of 27 commercial GLs were harvested at early, mid and late periods of the season. Fruit from each seasonal period were delivered to Massey University at two-week intervals with the first batch commencing 7 May, 2015 and the last batch arriving on 5 June, 2015. Each seasonal period consisted of nine unique growers arriving as 3 GLs on three different days during the same week. Each selected GL contained 3 trays of fruit (count 30), resulting in a total of 2430 fruit for all 27 GLs.

2.2.2. Segregation of fruit based on Vis-NIR measurements

At each delivery, the at-harvest Vis-NIR reflectance data of each fruit were obtained following the steps described in Section 2.1.2. Using this spectral data, the probability for each fruit to turn soft after 125 days of storage at 0ºC was estimated by the model developed using LogitBoost DS (Section 2.1). Fruit with estimated probability greater than 0.5 were classified as 'Soft' (high likelihood of failing the minimum firmness criterion); otherwise, fruit were classified as 'Good' (high likelihood of meeting the minimum firmness criterion).

For within GL segregation, ninety fruit from each GL arrived in three trays (30 fruit per tray) and were initially numbered. Fruit were then ranked based on descending probability and repacked into a new set of three trays, where the fruit from the same storage potential group were more likely to be sorted next to each other, while being separated from those from the other class. This method ensured an equal and maximum number of fruit in each tray. An alternative experimental method would have been to segregate fruit into trays by the class label of the fruit. However, this would have resulted in creating unbalanced trays because there was often not an even split between the two groups. Hence this method was avoided due to the introduced risk of influencing firmness by changing the water loss dynamics within the polylined tray that would occur with changes in product mass inside the tray.

For between GL segregation, no repacking was required because individual GLs were separated by default, as a result of storing fruit in units of three trays per GL (Fig.2). In this case the percentage of short-storing fruit for each GL was calculated based on the abovementioned probability estimated by the model. The 27 GLs were then ranked based on the calculated percent of short-storing fruit, with the top 9 lines being the longest storing fruit ($\leq$ 10%) and the remaining being 9 lines each of medium ($>$ 10% and $\leq$ 30%) and short storing fruit ($>$ 30%), respectively.

2.2.3. Coolstorage and destructive FF measurements

After segregation and repacking, fruit were immediately stored in a coolroom at 0 °C. During storage, ethylene concentration in the cool room was monitored and maintained below 5 nL L^{-1}. After 125 days of storage, FF measurements were performed on individual fruit using the methods described in Section 2.1.3., except that in this experiment the penetration speed was changed to 8 mm·s^{-1} due to a change in standard industrial procedure.

2.2.4. Data analysis

The potential for within GL segregation was assessed by determining if differences in FF values were obtained between the three different trays of fruit within the same GL as a result of the sorting and repacking applied. The effect of segregation on average FF within GL was analysed using the general linear model (GLM) in Minitab® (Version 16.1.0, Minitab Inc., Pennsylvania, USA). The number of soft fruit in different trays within GL was also analysed using a χ-square test in Minitab® to indicate evidence of potential benefit.

The performance of between GL segregation was assessed by comparing the predicted and actual storage performance of different lines. The percentage of accurate prediction was estimated for each of the short, medium and long storing GLs. The correlation between predicted and measured percent of soft fruit was established using Fitted Line Plot regression in Minitab®. Additionally, the proportions of soft fruit found in the segregated GL groups were compared using a χ-square test in Minitab®.

In addition, the percentage of fruit loss with and without segregation was calculated to indicate whether segregation prior to storage has the potential to reduce costs and improve profitability in the industry. For within GL segregation, reduction was estimated on the basis that fruit stored in the short- and medium-storing trays were to be shipped earlier whereas fruit in the long-storing trays were to be kept for later in the season. For between GL segregation, reduction assumed that predicted short- and medium-storing lines were to be distributed earlier whereas long-storing lines would be shipped later in the season.

3. Results and Discussion

3.1. Classification model for kiwifruit storage potential

3.1.1. Classification performance

Fig. 3 illustrates the at-harvest spectral data of kiwifruit from the training dataset grouped by their corresponding post-storage FF values. The preprocessed spectra showed several normalised absorption peaks and valleys associated with the physiochemical properties of the kiwifruit measured at harvest. This includes two valleys and a peak observed in the visible range (520 nm, 650 nm and 690 nm) typically associated with absorption bands of combined carotenoid and chlorophyll (520–530 nm; Merzlyak et al., 2003); anthocyanin pigment-sugar-protein complex (530–550 nm; Merzlyak et al., 2003; Liu et al., 2017), and chlorophyll (680 nm; McGlone and Kawano, 1998; 650 nm; Merzlyak et al., 2003). Absorption peaks around 950 nm, 1140 nm, 1390 nm, 1880 nm are associated with overtones and combination features of functional groups such as O-H (958 nm; McGlone and Kawano, 1998; 960 nm, 1440 nm and 1900 nm; Walsh et al., 2020a) and C-H bonds (1100 nm; Walsh et al., 2020a). The average spectra of kiwifruit between the two post-storage FF groups showed minimum differences at harvest, except for slightly lower absorption of the Soft group in the range of 400 – 530 nm, indicating a potential influence of at-harvest combined carotenoid and chlorophyll and/or anthocyanin concentrations on the storage potential of the same fruit. Lower absorption (i.e. pigment concentrations) seemed to correlate to poorer storage potential.

On the other hand, slightly higher absorption was observed at 780 nm and 1000 nm for the Soft fruit, suggesting a possible difference in water and carbohydrate contents between the two FF groups at the time of harvest, which is likely related to the maturity of the fruit. Famiani et al. (2012) found that higher initial DMC (hence lower water content) was associated with better firmness retention during storage. Similarly, Tombesi et al. (1993) found that the higher initial water content in shaded kiwifruit resulted in high transpiration during storage, which consequentially reduced cell turgor and FF during storage. This suggests that fruit that soften more quickly during storage are more likely to have higher water content at harvest. This is also evidenced by the higher absorptions observed at 780 nm and 1000 nm for the Soft fruit (Fig. 3).

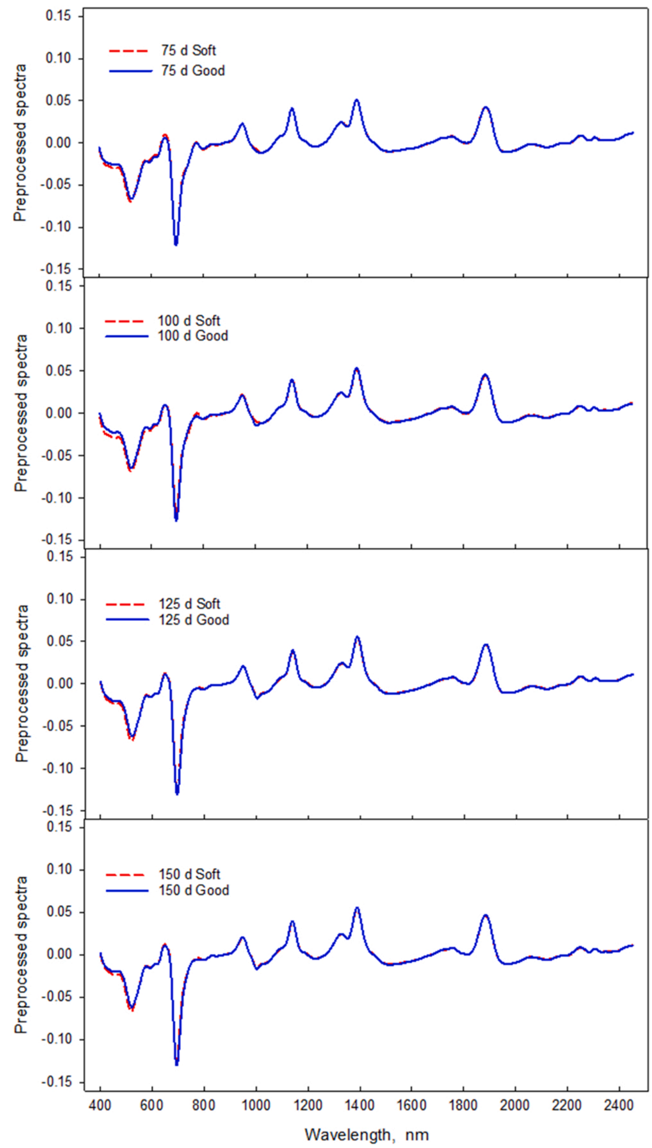

Fig. 3. Preprocessed at-harvest near infrared spectral data (absorbance, 1st derivative and normalisation) for 'Hayward' kiwifruit corresponding to post-storage flesh firmness (FF) stored at 75, 100, 125 and 150 days of cool storage at 0 °C. Solid blue line represents the average at-harvest spectral data for fruit that remained export quality, i.e. Good (FF $\geq$ 9.8 N) after storage. Dashed red line represents the average at-harvest spectral data for fruit that became soft and thus unexportable, i.e. Soft (FF $<$ 9.8 N) after storage.

The storage performance data appeared to be highly imbalanced, with the soft fruit class under-represented relative to the good fruit class in all data sets (Table 1). As the storage time increased the number/proportion of soft fruit generally increased and hence the data became relatively more balanced. Sample size seems to have affected predictive performance of the classification models, with better prediction accuracy observed for the Good group (> 78%) than the Soft group (< 54%) for all classifiers. In general prediction of soft fruit was better at 125 and 150 days (~ 40 – 50%) compared to that at 75 and 100 days (~ 20 – 40%; Table 2). This could also be related to the higher number of soft fruit (hence, more balanced data sets) found at those storage times. Results obtained for 125 and 150 days were comparable to the results found by Feng (2003) for the classification of healthy and soft-patched fruit (67% and 35% respectively), and those reported by Clark et al. (2004) for segregation between good and disordered fruit (~85% and

~59% respectively).

3.1.2. Classification algorithm comparison

Table 3 summarises the statistics for model evaluation and Garrett's ranking of model performance amongst classifiers. The top two ranking classifiers were SVM and LogitBoost DS regardless of storage time, with LogitBoost DS performing better at 125 and 150 days whereas SVM performing better at 75 and 100 days as also indicated by the AUC values at those storage times (Table 3). The SVM and LogitBoost had similar OA but SVM had higher Recall at 75, 100 and 150 days whereas LogitBoost DS had lower FN rates at all times (Table 3). In this case low FN rates are preferable because the proportion of soft fruit in the predicted good class should be as low as possible in order to ensure exportability and long storability of the predicted good batch. It is also important to note that when Garrett's Ranking Technique was applied, equal importance was assigned to all the performance metrics. As such the rankings could be modified by assigning different weightings to the selected metrics, should such requirements be needed for a particular case.

The 'No Free Lunch Theorem' (Wolpert and Macready, 1997) suggests that "any two optimisation algorithms are equivalent when their performance is averaged across all possible problems". There is no single classifier that had the best performance under all circumstances. Predictive performance amongst the classifiers varied at different storage times, and was dependent on the parameters selected for evaluation (Table 3). Therefore it is more sensible to say that there is no absolute best learning algorithm. The choice of the most suitable classifier is dependent on the nature of the data set and the criteria for discriminant analysis. In this case LogitBoost and SVM were the better classifiers because they performed well in cases that are more critical for the purposes of this study. Alternatively, ensemble methods which combine weighted individual decisions from a set of algorithms often can produce more accurate predictions than individual classifiers.

Nonetheless, the observed differences in performances amongst classifiers have been addressed and explained in many previous studies. Dettling and Bühlmann (2003) suggested that the lower error rates obtained with LogitBoost were because, unlike AdaBoost which uses an exponential function, LogitBoost uses the binomial log likelihood, which increases linearly rather than exponentially for strong negative margins. As such LogitBoost usually performs better on noisy data (Dettling and Bühlmann, 2003). This is important for the current study as the variation in firmness measurements by penetrometer (Li et al., 2017) could result in a large margin of error for sample grouping and hence, resulting in misclassification of fruit. Caruana and Niculescu-Mizil (2008) suggested that SVM had high error because the measurement of error interprets predictions as posterior probabilities but SVM is not designed to predict probabilities; the output of an SVM are just normalised distances to the decision boundary. In addition, Wu et al. (2010) suggested that LogitBoost DS was more capable of handling mixed data because, unlike SVM which relies on the Euclidean distance between two data points, the decision split at each stump branch does not rely on any particular distance measure between any pair of data points. Hence it should be more robust to outliers in both input and feature spaces.

Considering the lower FN rates and the industrial implications which benefit more from segregation of storability beyond 100 days, LogitBoost DS would be more suitable to be used to segregate potential soft fruit from the entire batch. Because of this, the model developed for 125 days after storage using LogitBoost DS was selected for external validation.

Table 2
Calibration of classification models to predict kiwifruit storage potential based on at-harvest Vis-NIR spectra data (original data) using 10-fold cross validation. Data represents classification of 2125 fruit from 4 trials in 2012 – 2013 at different storage times. Classification accuracy represents the fruit with accurately predicted FF groups (based on the minimum export criterion of 9.8 N) as a percent of the actual fruit belonging to that FF group at a specific time point.

Storage time (day)	Classification Accuracy (%)							
	Random Forest		Support Vector Machine		AdaBoost Decision Stumps		LogitBoost Decision Stumps	
	Soft	Good	Soft	Good	Soft	Good	Soft	Good
75	3	99	46	91	17	96	17	97
100	2	99	41	96	20	98	26	96
125	34	86	50	79	50	78	54	79
150	38	94	51	91	47	91	41	96

3.2. Segregation of kiwifruit storage potential using the developed model – an external validation

3.2.1. Within grower lines segregation

The average firmness amongst the three trays within the same GL differed as a result of pre-storage segregation (Table 4). Fruit from the first trays (highest probability of becoming soft at the end of storage) had lowest average firmness (12.56 N), whereas those from the third trays (lowest probability of becoming soft at the end of storage) had highest average firmness (14.03 N). The percent of soft fruit was highest in the first trays (30.7%), followed by the second (23.3%) and the third trays (21.2%). The χ-square analysis confirmed that the proportion of soft fruit was significantly different amongst the three trays ($p < 0.001$) as a result of ranking prior to storage. In addition, the effect of segregation within GL seemed to be more pronounced in the fruit harvested during early seasonal period (especially in G1 – G6) than those in late seasonal fruit (e.g. G22 – G27; Table 4). The effect also seemed to be more pronounced within growers with lower proportion of soft fruit (e.g. G1 – G6, G11, G21) but there were a few exceptions (e.g. G17).

Fruit with lower at-harvest DMC (i.e. higher at-harvest water content) have been reported to have poorer firmness retention during storage (Famiani et al., 2012; Tombesi et al., 1993). In this case, the at-harvest NIR reflectance signal could be attenuated due to stronger absorption in water absorption bands, masking any significant peaks over absorption bands by other chemicals such as pectin. As a result, it is possible that some information was lost during NIR data capture which would in turn affect model prediction. This may explain the poor prediction of GLs with higher proportion of soft fruit. However, the DMC range in the mentioned previous studies was between 13% and 16%, lower than the likely range of DMC for the fruit used in this study (usually $\approx$ 18% for 'Hayward'). Hence, it is also possible that the finding in these studies may not be directly relevant to the current study.

Softening of kiwifruit occurs in two or three phases depending on the maturity at harvest (Lallu et al., 1989). Kiwifruit that are harvested late in the season only go through the second and third stages; there is no lag phase during softening. Redgwell and Percy (1992) found that little pectin solubilisation was observed during the softening of kiwifruit from about 81–56 N (during the lag phase), whereas pectin became more soluble as fruit continued to soften below 56 N (the second and third phases). In addition, pectin undergoes a physical change from a rigid, solid-like state retained in the cell wall in firm fruit, to a more mobile, liquid-like state as fruit start the lag phase of softening (Schröder and Atkinson, 2006). It is possible that the structural features of cell wall components during the lag phase as influenced by the physical change of pectin played a more important role or provided more comprehensive information for model prediction. As a result, segregation within GL was more successful in those lines harvested during early seasonal period. However, it is also possible that this change is simply too insignificant to be detected by NIR signals owing to the low pectin concentrations (< 1%) in kiwifruit (McGlone and Kawano, 1998).

Statistically significant differences amongst trays suggest that prediction based on segregation was not simply a result of random variability. The model was able to identify a proportion of the actual short-storing fruit from the population. Repacking ensured that short-storing

Table 3
Ranking of classifier performance in predicting storability of kiwifruit after coolstorage using 10-fold cross-validation. SVM means support vector machines classification. DS means decision stumps. ROC denotes receiver operating characteristic curve and AUC denotes area under curve. MAE refers to mean absolute error. FN refers to false negative rate.

Storage time (days)	Rank	Classifier	Evaluation Parameter (Range)						
			Overall Accuracy (0 – 100%)	ROC AUC (0 – 1)	Kappa (0 – 1)	MAE (> 0)	FN (0 – 1)	Recall (0 – 1)	Precision (0 – 1)
75	1	SVM	86.567	0.830	0.3404	0.1459	0.087	0.457	0.381
	2	LogitBoost DS	88.358	0.778	0.1816	0.1243	0.033	0.171	0.375
	3	AdaBoost DS	88.060	0.800	0.1744	0.1258	0.037	0.171	0.353
	4	Random Forest	88.955	0.781	0.0305	0.1714	0.010	0.029	0.250
100	1	SVM	90.292	0.856	0.4120	0.1213	0.042	0.414	0.529
	2	LogitBoost DS	89.240	0.814	0.2791	0.1236	0.036	0.264	0.451
	3	AdaBoost DS	89.708	0.815	0.2340	0.1195	0.023	0.195	0.486
	4	Random Forest	89.591	0.857	0.0303	0.1534	0.005	0.023	0.333
125	1	LogitBoost DS	70.149	0.733	0.3389	0.3290	0.206	0.537	0.596
	2	SVM	69.697	0.720	0.2966	0.3561	0.215	0.504	0.570
	3	AdaBoost DS	67.463	0.727	0.2781	0.3385	0.224	0.496	0.556
	4	Random Forest	66.866	0.702	0.2127	0.4010	0.145	0.339	0.569
150	1	LogitBoost DS	83.833	0.859	0.4431	0.2268	0.041	0.414	0.743
	2	SVM	82.500	0.840	0.4563	0.1940	0.086	0.511	0.630
	3	Random Forest	81.833	0.858	0.3822	0.2469	0.058	0.383	0.654
	4	AdaBoost DS	81.200	0.840	0.4151	0.2019	0.086	0.466	0.608

Table 4
Average flesh firmness (N) and number of soft fruit amongst the three trays (T1 – T3; 30 fruit per tray) within a grower line after storage at 0 °C for 125 days as a result of pre-storage within grower line segregation.

Grower	Average firmness (N)			No. of soft fruit		
	FF T1	FF T2	FF T3	n T1	n T2	n T3
G1	13.44^{b}	14.52^{b}	17.07^{a}	4	3	0
G2	12.65^{b}	14.32^{b}	18.44^{a}	7	3	2
G3	14.32^{b}	17.66^{a}	19.72^{a}	2	0	0
G4	14.52^{b}	16.97^{ab}	19.03^{a}	1	0	1
G5	15.60^{b}	17.85^{ab}	19.52^{a}	3	0	1
G6	14.52^{b}	17.56^{a}	18.64^{a}	1	0	1
G7	13.34^{a}	13.93^{a}	14.13^{a}	9	5	6
G8	8.24^{a}	9.81^{a}	8.83^{a}	22	19	17
G9	14.81^{a}	13.83^{a}	15.50^{a}	2	4	3
G10	11.87^{a}	11.97^{a}	12.75^{a}	6	5	6
G11	15.70^{b}	17.27^{ab}	18.15^{a}	2	0	0
G12	12.95^{a}	14.42^{a}	15.50^{a}	8	3	6
G13	12.75^{a}	13.15^{a}	14.03^{a}	5	1	3
G14	18.44^{a}	19.91^{a}	18.74^{a}	2	0	0
G15	10.59^{a}	10.79^{a}	10.20^{a}	13	10	11
G16	15.50^{a}	17.46^{a}	17.56^{a}	3	0	0
G17	10.20^{b}	10.69^{b}	14.42^{a}	14	12	4
G18	11.18^{a}	12.26^{a}	10.99^{a}	10	8	12
G19	9.61^{a}	10.59^{a}	10.01^{a}	15	13	14
G20	9.22^{a}	9.81^{a}	9.81^{a}	17	14	13
G21	12.46^{ab}	11.38^{b}	13.15^{a}	2	6	2
G22	10.69^{a}	10.40^{a}	11.28^{a}	10	8	5
G23	9.03^{a}	9.52^{a}	9.71^{a}	25	17	18
G24	9.81^{a}	10.40^{a}	10.79^{a}	14	11	9
G25	9.81^{a}	10.10^{a}	10.40^{a}	17	13	10
G26	10.30^{a}	11.48^{a}	11.28^{a}	12	8	6
G27	7.06^{a}	6.77^{a}	8.14^{a}	23	26	22
Total	12.56^{c}	13.15^{b}	14.03^{a}	249	189	172

fruit were separated from the remaining population, reducing the exposure to ethylene for those long-storing fruit.

3.2.2. Between grower line segregation

The model predicted that the majority of the GLs harvested during early seasonal period (G1 – G9) would have medium or short storability; these lines would likely to have 17.8 – 54.4% (16 – 49 fruit) of the population turning soft (Fig. 4a). However, G6 was predicted to be an exception: only 8.9% (8 fruit) would be turning soft at the end of storage (Fig. 4a). The late harvested fruit (G19 – G27) were predicted to perform similarly to those harvested during early seasonal period, with most lines having medium or short storability (18.9 – 47.8% failure; 17 – 43 soft fruit) except for G24 being long-storing (4.4% failure; 4 soft fruit; Fig. 4a). For GLs harvested during mid seasonal period (G10 – G18), the model predicted that most lines were to have long storage with 3.3 – 10% failure (3 – 9 fruit). The predicted percent of soft fruit was lowest for mid seasonal fruit but was higher in both early and late seasonal fruit (Fig. 4a).

Post-storage firmness measurements showed that there was a large variation in storage potential between GLs (Fig. 4b). In general, fruit harvested during early and mid-seasonal periods (G1 – G18) had lower proportion of soft fruit and higher average firmness after 125 days of storage compared to those harvested during late seasonal period (G19 – G27). The number/percent of actual soft fruit also generally increased from early to late seasonal periods (Table 4; Fig. 4a). Most early harvested lines had long to medium storability (2.2 – 22.2% failure; 2 – 20 soft fruit) except for G8 which had 58 soft fruit (64.4% failure). Similarly, most of the mid harvested lines had long to medium storability (2.2 – 18.9% failure; 2 – 17 soft fruit) except for G15, G17 and G18 (37.8%, 33.3% and 33.3% soft fruit, respectively). Although the predicted and the measured number of soft fruit was somewhat similar (602 and 610, respectively), the model overestimated the percent of soft fruit in early harvested lines (e.g. GLs 3, 7 and 9), but underestimated that in the mid harvested lines (e.g. GLs 15, 18; Fig. 4a).

Table 5 displays the classification accuracy for segregation between GLs for storability. For each of the short, medium and long storage group, 4, 3 and 4 out of 9 lines were classified accurately respectively. The percentage of actual short, medium and long storage lines in the predicted short, medium and long storage lines was 44.4%, 33.3% and 44.4%, compared to 37.0%, 25.9% and 37.0% respectively by chance (Table 5). The overall accuracy (40.7%) is not as good compared to the 52.8% accuracy achieved by Feng (2003) for the segregation of 36 GLs using CDA based on several at-harvest attributes including solublised DMC, harvest date, fruit lightness and mineral content. This is probably because in Feng's study, results were obtained from internal validation data set and hence were not subjected to errors as a result of unknown variability from an independent new data set.

The prediction in this study showed inferior performance of the

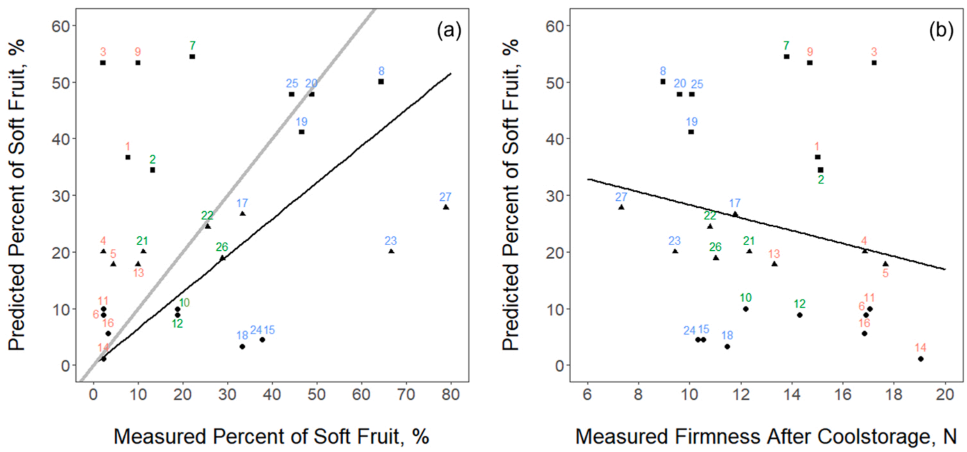

Fig. 4. Relationship of (a) predicted versus measured proportion of soft fruit, (b) predicted proportion of soft fruit versus post-storage firmness measurements for 27 grower lines sourced in2015. Circle, triangle and square shapes represent grower lines with predicted long (9), medium (9) and short (9) storability, respectively. Red, green and blue labels are grower line numbers with actual long (10), medium (7) and short (10)storability, respectively. Solid black lines are fitted linear regression lines based on all data points whereas the solid grey line represents the perfect prediction line (i.e. y = x).

Table 5
Accuracy of classification model developed using LogitBoost decision stumps based on segregation into three groups: short (> 30% soft fruit), medium (10 – 30% soft fruit) and long (≤ 10% soft fruit) storability between 27 grower lines.

	Predicted number (percentage) of grower line				Average FF (N)
Actual number	Short	Medium	Long	Total	
Short	**4 (44.4%)**	3 (33.3%)	3 (33.3%)	10 (37.0%)	9.96
Medium	2 (22.2%)	**3 (33.3%)**	2 (22.2%)	7 (25.9%)	12.79
Long	3 (33.3%)	3 (33.3%)	**4 (44.4%)**	10 (37.0%)	16.46
Total	9	9	9	27	–
Average FF (N)	12.74	12.27	14.29	–	

medium storing lines, in accordance with the poorer prediction of medium storage lines (33.3 – 46.7%) than the short or long storage lines (60.0 – 80.0%) as observed by Feng (2003). This is not surprising as fruit that were assigned to the medium-storing group had estimated probability around 0.5 (at the boundary of the binary threshold for firmness), indicating reduced confidence in the model's predictability leading to higher chances of errors in storability prediction. The average FF for the predicted long, medium and short storage groups was 14.29 N, 12.27 N and 12.74 N respectively, compared to 16.46 N, 12.79 N and 9.96 N respectively for the actual groups (Table 5). Both predicted and actual long-storing groups had the highest average FF. However, the lowest FF was found in the medium-storing group based on prediction.

3.2.3. Soft fruit reduction

In the external validation data set, 25.1% (610 fruit) of the total population were soft after 125 days of coolstorage at 0 °C (Fig. 5a and Fig. 5b). At the time of harvest the model accurately predicted 196 out of 610 soft fruit and 1418 out of 1820 good fruit (Table 6). The TP and TN rates were 32.1% and 77.9% respectively, compared to 24.6% and 75.4% respectively by chance. The FN rate was 22.6% (414 out of 1832) which indicate the actual fruit loss in the predicted good batch (Table 6).

The predictive accuracy for 'Soft' (32.1%) and 'Good' (77.9%) was slightly lower than the 54% and 79% obtained in the calibration model for 'Soft' and 'Good', respectively (Table 2). The true positive rate (32.1%) was comparable to the 17 – 47% for kiwifruit that developed soft patches found by Feng (2003) but was lower compared to the 46 – 63% for kiwifruit that developed disorders found by Clark et al. (2004). The false positive rate (22%; 402 fruit out of 1820) was lower than the 33 – 86% for healthy fruit found by Feng (2003), but was higher than those obtained in Clark et al. (2004) for good fruit (8 – 18%).

According to Hair et al. (2006), the classification accuracy should be at least one-fourth (25%) greater than that achieved by chance, in order to justify the significance of improvement using a discriminant model. In this study, the classification accuracy for soft fruit (32.1%) was more than 25% greater than that achieved by chance (24.6%). However, the classification accuracy for good fruit was 77.9%, less than the 94.3% (75.4% × 1.25) required to justify a significant segregation. This suggests that in general, classification for soft fruit resulted in more significant improvement compared to that for good fruit, despite a lower % classification accuracy.

For segregation within grower lines, it was intended that segregation followed by ranking and repacking of fruit would reduce final soft fruit proportion by allowing earlier shipping of fruit stored in the first and second trays while keeping those in the third trays for later in the season. Post-storage firmness results showed that 438 out of 1620 (27.0%) early-shipment fruit (first and second trays) were soft, whereas 638 out of 810 (78.8%) late-shipment fruit (fruit stored in the third trays) were sound (Fig. 5a). Segregation resulted in two populations: early shipment with a slightly higher proportion of short-storing fruit than the whole population and late shipment with a higher proportion of long-storing fruit. The proportion of soft fruit reduced by 15.5% from 25.1% (shipment without segregation) to 21.2% (shipment based on segregation). The proportion of soft fruit was significantly different between the two segregated populations ($\chi^2 = 9.670$; p = 0.002). However, the reduction of soft fruit (15.5%) was not significant according to the minimum criterion (25%) as defined by Hair et al. (2006).

For segregation between grower lines, it was intended that the predicted short- and medium-storing GLs would be shipped earlier in the season while the predicted long-storing lines would be kept for later sales. Post-storage firmness results showed that 469 out of 1620 (29.0%) fruit in the early shipment lines were soft, whereas 669 out of 810 (82.6%) fruit in the late shipment were sound (Fig. 5b). The proportion of soft fruit reduced by 30.7% from 25.1% (shipment without segregation) to 17.4% (shipment based on grower line segregation). The proportion of soft fruit in the two segregated populations was significantly different ($\chi^2 = 38.270$; $p < 0.001$). The reduction of soft fruit (30.7%) was significant (> 25%), suggesting the potential industrial applicability of the developed segregation system on current packing lines.

The performance of the model in this validation trial was comparable to that during model calibration (Section 3.1.1.), suggesting good repeatability of the model. This is possibly attributed to the wide variety of factors (growth conditions, maturity, grower lines and seasonal variation) included in the training samples in the calibration data set. The slight reduction in ability to predict between model calibration and validation is to be expected. The purpose of validation is to ensure that results obtained in the calibration model are not specific to only the calibration sample, but could also be generalised to an independent population (Hair et al., 2006). However, multivariate models in food processes can take years to build and/or improve up to a desirable level.

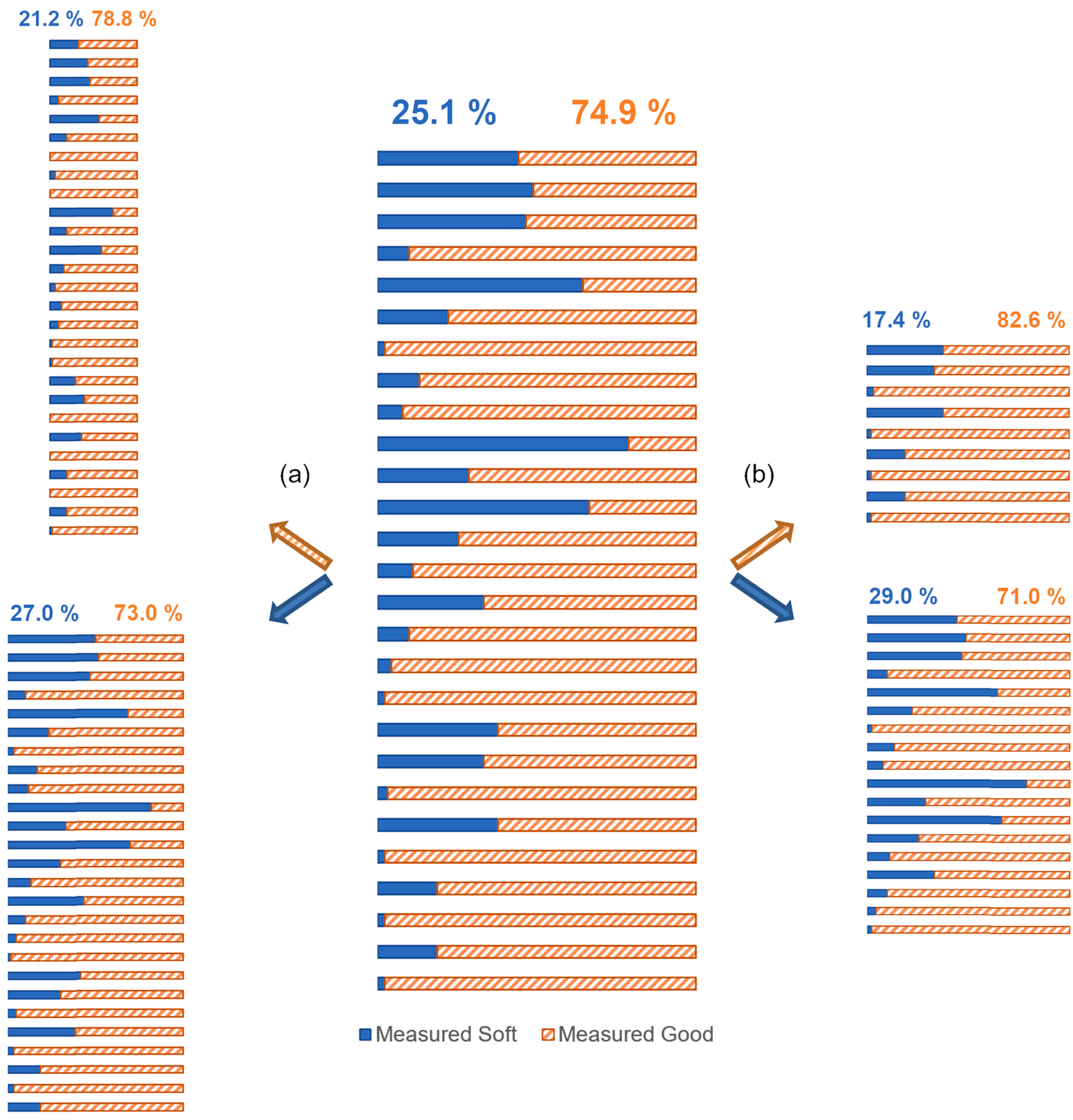

Fig. 5. Improvement in storage performance based on (a) segregation of trays with shorter storability within grower line at the time of harvest using Vis-NIR spectroscopy, and (b) segregation between grower lines with short and medium storability using data collected in 2015. Hatched orange and solid blue bars represent the actual soft (FF $<$ 9.8 N) and good (FF $\geq$ 9.8 N) fruit in whole validation population. The solid blue arrows lead to the population of fruit stored either in the first and second trays of a grower line, or short- and medium-storing grower lines. The hatched orange arrows lead to the population of fruit stored in the third trays of a grower line, or long-storing grower lines.

The lack of selectivity in the multivariate NIR signal and the time span required for model building can be problematic when unwanted sources of variation are present in the data (Sileoni et al., 2011). This is especially true for natural products such as food raw materials (e.g. freshly harvested fruit) in open biological processes where the composition cannot be exactly predetermined. In this study, the predictive accuracy could have been improved by recalibration of the model by adding a set of NIR spectral and fruit storage data obtained from the validation trial. However this is impractical because prediction needs to take place prior to cool storage.

For between grower line segregation, the fact that the model overestimated the proportion of soft fruit for GLs harvested during early seasonal period suggests the lack of reliability to rank GLs according to their harvesting time. It is recommended that, for future improvement, harvesting time (e.g. ISO days) or harvesting seasonal periods (e.g. early, mid or late) should be included in the calibration model as an input variable, so as to improve the prediction accuracy of the model. Nonetheless, grower line remains the most likely contributor to the source of firmness variability in the supply chain because the status of fruit at harvest is influenced by orchard operations, microclimates and local weather and soil conditions for a particular season.

Table 6
Confusion matrix for external validation using LogitBoost decision stumps classification model.

	Predicted		
Actual	Soft	Good	Total
Soft	196 (32.1%)	**414** **(67.9%)**	610 (25.1%)
Good	402 (22.1%)	1418 (77.9%)	1820 (74.9%)
Total	598 (24.6%)	1832 (75.4%)	2430

4. Conclusions

The storage potential of individual fruit can be predicted based on its likelihood to become unacceptable for export (FF < 1 kg_f or 9.8 N) after coolstorage using at-harvest Vis-NIR spectral data and the post-storage FF collected at various time points. Segregation of individual fruit within a grower line using the developed classification model prior to storage resulted in significantly different post-storage FF means amongst the segregated trays, with lowest FF found in the population having highest probability for the soft group. This would contribute to a 15.5% fruit loss reduction if the predicted soft group is sold early in the season. The storage potential between grower lines could also be predicted, achieving a 30.7% fruit loss reduction should the segregated short and medium storing lines be shipped earlier in the season.

This work demonstrates the value of segregation strategies to reduce fruit loss / food waste in postharvest horticultural systems. External validation is a crucial process which provide a realistic assessment of industrial applicability for NIR-based sorting technologies, and therefore should not be overlooked in practical applications of segregation and decision-making strategies from NIR data.

CRediT authorship contribution statement

Mo Li: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Reddy Pullanagari:** Methodology, Resources, Writing – review & editing, Supervision. **Ian Yule:** Resources, Writing – review & editing, Supervision. **Andrew East:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work is an output of the Food Structure Design Theme of the Fonterra / Zespri® Primary Growth Partnership research project funded by the Ministry of Primary Industries, New Zealand and supported by Zespri® International Ltd.

References

Aitken, A., Warrington, I., 2019. FreshFacts - New Zealand horticulture. Data retrieved on 21st Dec, 2020, from https://www.freshfacts.co.nz/files/freshfacts-2018.pdf.
Anne-Michelle, N., Mousumi, B., 2007. Machine learning methods for cancer diagnosis and prognostication. In: Khattree, R., Naik, D.N. (Eds.), Computational Methods in Biomedical Research. CRC Press, pp. 77–101.
Beever, D.J., Hopkirk, G., 1990. Fruit development and fruit physiology. In: Warrington, I.J., Weston, G.C. (Eds.), Kiwifruit: science and management. New Zealand Society for Horticultural Science, Auckland, pp. 97–126.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Burdon, J., Wohlers, M., Pidakala, P., Laurie, T., Punter, M., Billing, D., 2014. The potential for commonly measured at-harvest fruit characteristics to predict chilling susceptibility of 'Hort16A' kiwifruit. Postharvest Biol. Technol. 94, 41–48. https://doi.org/10.1016/j.postharvbio.2014.03.005.
Buxton, K.N., 2005. Preharvest practices affecting postharvest quality of 'Hayward' kiwifruit: a thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy. Plant Physiology and Horticultural Science. Massey University,, Palmerston North, New Zealand. ⟨http://hdl.handle.net/10179/1630⟩.
Clark, C.J., McGlone, V.A., De Silva, H.N., Manning, M.A., Burdon, J., Mowat, A.D., 2004. Prediction of storage disorders of kiwifruit (*Actinidia chinensis*) based on visible-NIR spectral characteristics at harvest. Postharvest Biol. Technol. 32, 147–158. https://doi.org/10.1016/j.postharvbio.2003.11.004.
Costa, G., Andreotti, C., Miserocchi, O., Noferini, M., Smith, G., 1999. Near-infrared (NIR) methods to determine kiwifruit field harvest date and maturity parameters in cool store. ActaHortic 498, 231–238. https://doi.org/10.17660/ActaHortic.1999.498.26.
Dettling, M., Bühlmann, P., 2003. Boosting for tumor classification with gene expression data. Bioinformatics 19, 1061–1069. https://doi.org/10.1093/bioinformatics/btf867.
East, A.R., Jabbar, A., Heyes, J.A., 2013. Approaches to prediction of storage out-turn for units of fresh produce. ActaHortic 1012, 1303–1309. https://doi.org/10.17660/ActaHortic.2013.1012.175.
Famiani, F., Baldicchi, A., Farinelli, D., Cruz-Castillo, J.G., Marocchi, F., Mastroleo, M., Moscatello, S., Proietti, S., Battistelli, A., 2012. Yield affects qualitative kiwifruit characteristics and dry matter content may be an indicator of both quality and storability. Sci. Hortic. 146, 124–130. https://doi.org/10.1016/j.scienta.2012.08.009.
Feng, J., 2003. Segregation of 'Hayward' kiwifruit for storage potential: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy. Massey University, Palmerston North, New Zealand. http://hdl.handle.net/10179/2376.
Feng, J., Stanley, J., Othman, M., Woolf, A., Kosasih, M., Olsson, S., Clare, G., Cooper, N., Wang, X., 2013. Segregation of apricots for storage potential using non-destructive technologies. Postharvest Biol. Technol. 86, 17–22. https://doi.org/10.1016/j.postharvbio.2013.06.015.
Feng, J., Wohlers, M., Olsson, S., White, A., McGlone, V.A., Seelye, R., Tanner, D., Bollen, F., 2014. Comparison between an acoustic firmness sensor and a near-infrared spectrometer in segregation of kiwifruit for storage potential. Acta Hortic. 1119, 279–288. https://doi.org/10.17660/ActaHortic.2016.1119.39.
Freund, Y., Schapire, R.E., 1996. Experiments with a new boosting algorithm, Proceedings of the 13th International Conference on Machine Learning, Bari, Italy, pp. 148–156.
Friedman, J., Hastie, T., Tibshirani, R., 2000. Additive logistic regression: a statistical view of boosting. Ann. Stat. 28, 337–407.
Garrett, G., 2002. Our method explained. US News and World Report, 2003 edition.,. Best Graduate Schools,, Washington, D.C, pp. 34–35.
Gorunescu, F., 2011. Classification Performance Evaluation. In: Kacprzyk, J., Jain, L.C. (Eds.), Data mining: concepts, models and techniques. Springer-Verlag, Berlin, pp. 319–331.
Gunn, S.R., 1998. Support vector machines for classification and regression, Technical report. Faculty of Engineering, Science and Mathematics, School of Electronics and Computer Science, University of Southampton.
Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatam, R.L., 2006. Multiple Distriminant Analysis and Logistic Regression. In: Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatam, R.L. (Eds.), Multivariate Data Analysis. Pearson Prentice Hall, U.S.A, pp. 269–382.
Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The WEKA data mining software: An update. SIGKDD Explor. 11, 10–18. https://doi.org/10.1145/1656274.1656278.
Ivanciuc, O., 2007. Applications of support vector machines in chemistry. In: Lipkowitz, K.B., Cundari, T.R. (Eds.), Reviews in Computational Chemistry. John Wiley & Sons, Inc, Hoboken, New Jersey, pp. 291–400.
Jabbar, A., 2014. Accelerated Fruit Libraries to Predict Storage Potential of 'Hayward' Kiwifruit Grower Lines: A Thesis Presented in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Food Technology at Massey University, New Zealand.
Jabbar, A., East, A., 2016. Quantifying the ethylene induced softening and low temperature breakdown of 'Hayward' kiwifruit in storage. Postharvest Biol. Technol. 113, 87–94. https://doi.org/10.1016/j.postharvbio.2015.11.002.
Jung, J.J., 2009. Trustworthy knowledge diffusion model based on risk discovery on peer-to-peer networks. Expert Syst. Appl. 36, 7123–7128.
Kotsiantis, S.B., Kanellopoulos, D., Pintelas, P.E., 2006. Local boosting of decision stumps for regression and classification problems. J. Comput. 1, 30–37 http://dx.doi.org/10.1.1.90.4420.
Lallu, N., Searle, A.N., Macrae, E.A., 1989. An investigation of ripening and handling strategies for early season kiwifruit (*Actinidia deliciosa* cv Hayward. J. Sci. Food Agric. 47, 387–400. https://doi.org/10.1002/jsfa.2740470402.
Lee, J.S., Kim, S.C., Seong, K.C., Kim, C.H., Um, Y.C., Lee, S.K., 2012. Quality prediction of kiwifruit based on near infrared spectroscopy. Korean J. Hortic. Sci. 30, 709–717. https://doi.org/10.7235/hort.2012.12139.
Liu, Y., Sun, Y., Xie, A., et al., 2017. Potential of Hyperspectral Imaging for Rapid Prediction of Anthocyanin Content of Purple-Fleshed Sweet Potato Slices During Drying Process. Food Anal. Methods 10, 3836–3846. https://doi.org/10.1007/s12161-017-0950-y.
Li, M., Pullanagari, R.R., Pranamornkith, T., Yule, I.J., East, A.R., 2017. Quantitative prediction of post storage 'Hayward' kiwifruit attributes using at harvest Vis-NIR

spectroscopy. J. Food Eng. 202, 46–55. https://doi.org/10.1016/j.jfoodeng.2017.01.002.
McGlone, V.A., Kawano, S., 1998. Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy. Postharvest Biol. Technol. 13, 131–141. https://doi.org/10.1016/S0925-5214(98)00007-6.
Merzlyak, M.N., Solovchenko, A.E., Gitelson, A.A., 2003. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. Postharvest Biol. Technol. 27, 197–211. https://doi.org/10.1016/S0925-5214(02)00066-2.
Moghimi, A., Aghkhani, M.H., Sazgarnia, A., Sarmad, M., 2010. Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. Biosyst. Eng. 106, 295–302. https://doi.org/10.1016/j.biosystemseng.2010.04.002.
Mohri, M., Rostamizadeh, A., Talwalkar, A., 2012. Support vector machines. Foundations of machine learning. The MIT Press,, Massachusetts, USA, pp. 63–84.
Nguyen, H.-N., Vu, T.-N., Ohn, S.-Y., Park, Y.-M., Han, M.Y., Kim, C.W., 2006. Feature elimination approach based on random forest for cancer diagnosis. MICAI 2006: Advances in Artificial Intelligence. Springer,, pp. 532–542.
Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007a. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. Postharvest Biol. Technol. 46, 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024.
Caruana, R., Niculescu-Mizil, A. 2008. Getting the most out of your data: multitask bayesian network structure learning, predicting good probabilities and ensemble selection. PhD thesis. Cornell University, USA.
Osborne, S., Künnemeyer, R., Jordan, R., 1999. A low-cost system for the grading of kiwifruit. J. Near Infrared Spectrosc. 7, 9–15. https://doi.org/10.1255/jnirs.229.
Pachghare, V., Kulkarni, P., 2011. Pattern based network security using decision trees and support vector machine, Proceedings of the 3rd International Conference on Electronics Computer Technology (ICECT). IEEE, Kanyakumari, India, pp. 254–257.
Patterson, K.J., Currie, M.B., 2011. Optimising kiwifruit vine performance for high productivity and superior fruit taste. Acta Hortic. 913, 257–268. https://doi.org/10.17660/ActaHortic.2011.913.33.
Platt, J., 1999. Fast training of support vector machines using sequential minimal optimization. In: Schölkopf, B., Smola, A.J. (Eds.), Advances in Kernel Methods - Support Vector Learning. The MIT Press, Massachusetts, USA, pp. 41–65.
Quinlan, J.R., 1986. Induction of Decision Trees. Mach. Learn. 1, 81–106. https://doi.org/10.1023/a:1022643204877.
Redgwell, R., Percy, A., 1992. Cell wall changes during on-vine softening of kiwifruit. N. Z. J. Crop Horticult. Sci. 20, 453–456. https://doi.org/10.1080/01140671.1992.10418064.
Saeys, W., Nguyen Do Trong, N., Van Beers, R., Nicolaï, B.M., 2019. Multivariate calibration of spectroscopic sensors for postharvest quality evaluation: a review. Postharvest Biol. Technol. 158, 110981 https://doi.org/10.1016/j.postharvbio.2019.110981.
Samarakoon, H.C., 2013. Ethylene flux in postharvest kiwifruit systems: a thesis presented in partial fulfilment of the requirements for the degree of Master of Philosophy. Massey University, Palmerston North, New Zealand.
Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627–1639. https://doi.org/10.1021/ac60214a047.
Schaare, P.N., Fraser, D.G., 2000. Comparison of reflectance, interactance and transmission modes of visible-near infrared spectroscopy for measuring internal properties of kiwifruit (*Actinidia chinensis*). Postharvest Biol. Technol. 20, 175–184. https://doi.org/10.1016/S0925-5214(00)00130-7.
Schröder, R., Atkinson, R.G., 2006. Kiwifruit cell walls: towards an understanding of softening? New Zealand. J. Forest. Sci. 36, 112.
Sileoni, V., van den Berg, F., Marconi, O., Perretti, G., Fantozzi, P., 2011. Internal and external validation strategies for the evaluation of long-term effects in NIR calibration models. J. Agric. Food Chem. 59, 1541–1547. https://doi.org/10.1021/jf104439x.
Sun, D.W., 2009. Infrared spectroscopy for food quality analysis and control, 1st ed. Academic Press, Boston.
Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random forest: a classification and regression tool for compound classification and QSAR modeling. J. Chem. Info. Comput. Sci. 43, 1947–1958. https://doi.org/10.1021/ci034160g.
Tanner, D., Kay, C., Mowat, A.D., 2012. Segregation technologies - Are they ready? New Zealand Kiwifruit. Journal, July/August 15–18.
Tombesi, A., Antognozzi, E., Palliotti, A., 1993. Influence of light exposure on characteristics and storage life of kiwifruit. N.Z.J. Crop Horticult. Sci. 21, 85–90. https://doi.org/10.1080/01140671.1993.9513750.
Walsh, K.B., Blasco, J., Zude-Sasse, M., Sun, X., 2020a. Visible-NIR 'point' spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. Postharvest Biol. Technol. 168, 111246 https://doi.org/10.1016/j.postharvbio.2020.111246.
Walsh, K.B., McGlone, V.A., Han, D.H., 2020b. The uses of near infra-red spectroscopy in postharvest decision support: a review. Postharvest Biol. Technol. 163, 111139 https://doi.org/10.1016/j.postharvbio.2020.111139.
Williams, P.C., Norris, K.H., 1987. Near-infrared technology: in the agricultural and food industries, 1st ed. American Association of Cereal Chemists, Inc.
Witten, I.H., Frank, E., Hall, M.A., 2011. Data mining: practical machine learning tools and techniques. Morgan Kaufmann Publishers, Burlington, Massachusetts.
Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. IEEE Trans. Evolut. Comput. 1, 67–82.
Wu, J., Roy, J., Stewart, W.F., 2010. Prediction modeling using EHR data: Challenges, strategies, and a comparison of machine learning approaches. Med. Care 48, S106–S113. https://doi.org/10.1097/MLR.0b013e3181de9e17.
Zeaiter, M., Roger, J.M., Bellon-Maurel, V., 2005. Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods. TrAC Trends Anal. Chem. 24, 437–445. https://doi.org/10.1016/j.trac.2004.11.023.
Zude, M., Herold, B., Roger, J.-M., Bellon-Maurel, V., Landahl, S., 2006. Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life. J. Food Eng. 77, 254–260. https://doi.org/10.1016/j.jfoodeng.2005.06.027.