Journal of Food Engineering 131 (2014) 7–17

Contents lists available at ScienceDirect

Journal of Food Engineering

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

Modelling the relationship between peel colour and the quality of fresh mango fruit using Random Forests

Shinji Fukuda [a], Eriko Yasunaga [b,*], Marcus Nagle [c], Kozue Yuge [a], Vicha Sardsud [d], Wolfram Spreer [c,e], Joachim Müller [c]

[a] Faculty of Agriculture, Kyushu University, Hakozaki 6-10-1, Fukuoka 812-8581, Japan
[b] Institute for Sustainable Agro-ecosystem Services, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Midori Machi 1-1-1, Nishi Tokyo, Tokyo 188-0002, Japan
[c] Institute of Agricultural Engineering, Universität Hohenheim, Garbenstrasse 9, Stuttgart 70599, Germany
[d] Postharvest Technology Research Institute, Chiang Mai University, Chiang Mai 50200, Thailand
[e] Science and Technology Research Institute, Chiang Mai University, 239 Huay Kaew Road, Chiang Mai 50200, Thailand

ARTICLE INFO

Article history:
Received 26 August 2013
Received in revised form 2 December 2013
Accepted 10 January 2014
Available online 21 January 2014

Keywords:
Data mining
Non-destructive assessment
Postharvest ripening
Fruit quality
Peel colour
Mangifera indica L

ABSTRACT

Mango (*Mangifera indica* L.) is one of the major tropical fruits exported through long supply chains to export markets. Production of high quality fruits and monitoring postharvest changes during storage and transport are thus primary concerns for exporters to ensure the premium value of fresh mango fruit after distribution. This study aims to demonstrate the applicability of Random Forests (RF) for estimating the internal qualities of mango based on peel colour. Two cultivars, namely Nam Dokmai and Irwin, having different fruit properties and grown in intensively managed orchards in Thailand and Japan, respectively, were used in this study. Postharvest changes in peel colour and fruit quality were observed under three storage conditions with respect to temperature. RF models were applied to establish a relationship between peel colour and fruit quality, and then tested the applicability based on model accuracy and variable importance computed by the RF. Specifically, this work demonstrates how the variable importance can be used to interpret the model results. The high accuracy and the information retrieved by the RF models suggest the applicability and practicality as a non-destructive assessment method for the quality of fresh mango fruit.

1. Introduction

Mango (*Mangifera indica* L.) is one of the major tropical fruits, favoured for its taste, colour, texture and nutritional value. These fruit quality parameters depend strongly on both production (González et al., 2004; Spreer et al., 2007, 2009; Nagle et al., 2010) and postharvest management (Mahayothee et al., 2007; Kienzle et al., 2011, 2012). Mango exhibits climacteric behaviour, characterised by decreasing fruit respiration during development leading to a minimum, followed by a rise in respiration levels until full ripeness. Climacteric fruits are commonly harvested directly after the pre-climacteric minimum, meaning mature but not ripe for consumption, after which the fruit then undergoes post-harvest ripening during the climacteric rise (Grierson, 2002). The time of harvest influences the magnitude of the climacteric curve, and therefore, the final product quality (Seymour et al., 1990). For example, fruits harvested too early do not fully realise the desired ripening changes and a late harvest will lead to reduced shelf life and off-flavour (Medlicott et al., 1988; Lalel et al., 2003). In this regard, harvest maturity should be primarily considered for a better distribution management. Several methods are known for determining the optimum harvest time of mango that require a set of maturity-related physiological or quality attributes. The mango industry commonly uses destructive methods such as determination of flesh firmness and soluble solids content for assessing harvest maturity. However, harvest decisions based on fruit sampling are flawed because of high inconsistency between mango varieties and high variability within trees and orchards (Herold et al., 2005). Furthermore, access to technological resources, specialised equipment and expertise are needed. Most producers are unable to accurately determine the best time for harvest and remain using unreliable methods with no real standardization. Thus, a large potential exists for rapid, accurate, non-destructive sensor technology for predicting interior fruit qualities for determination of

* Corresponding author. Tel.: +81 42 463 1687; fax: +81 42 464 4391.
E-mail addresses: shinji-fkd@agr.kyushu-u.ac.jp (S. Fukuda), erikoy@isas.a.u-tokyo.ac.jp (E. Yasunaga), marcus.nagle@uni-hohenheim.de (M. Nagle), yuge@bpes.kyushu-u.ac.jp (K. Yuge), vicha_cmu@hotmail.com (V. Sardsud), wolfram.speer@uni-hohenheim.de (W. Spreer), joachim.mueller@uni-hohenheim.de (J. Müller).

http://dx.doi.org/10.1016/j.jfoodeng.2014.01.007
0260-8774/

optimal harvest time. Such sensors are useful because they allow measurement of every fruit, since they are non-destructive, and can be repeated while leaving the fruit on the tree until it is mature.

Postharvest ripening is a key process affecting the fruit quality during distribution, which is important especially for long supply chains. Changes in fruit properties during ripening are strongly affected by storage and transport conditions, such as temperature, humidity and atmospheric composition. The ability to monitor fruit quality changes during the postharvest handling chain, especially by non-destructive assessment methods, can help to ensure a premium product after distribution. Peel colour, paramount with respect to consumer acceptance, has been found to be one of the major fruit quality indicators for mangoes (Saranwong et al., 2004; Vásquez-Caicedo et al., 2005). However, Kienzle et al. (2011, 2012) reported that postharvest maturity stage is difficult to specify with respect to peel colour based on a 7-day monitoring period of comprehensive fruit quality parameters. Thus, it is worth investigating the relationships between fruit quality and peel colour parameters based on longer monitoring periods. For this purpose, predictive modelling methods, such as machine learning, can be applied. Such predictive models allow for establishing peel-colour-based systems for harvest maturity specification as well as automatic sorting systems that can, for instance, be used for sorting of mango fruit suitable either for domestic or export markets.

Quality prediction is one of the major interests in postharvest engineering (Hertog et al., 2011), in which machine learning methods, including artificial neural networks (ANNs), support vector machines (SVMs) and classification and regression trees (CARTs), can be useful as data-driven modelling approaches. For instance, ANNs have been used to detect chilling injury in apple based on hyperspectral imaging (ElMasry et al., 2009) and to predict a vase life of cut roses (In et al., 2009). Gomez-Sanchis et al. (2012) applied ANNs and CARTs for rottenness detection of citrus fruits. SVMs have been increasingly applied, for instance, to detect browning degree of mango fruit (Zheng and Lu, 2012), or to detect bruising of red bayberry (Lu et al., 2011) and apples (Baranowski et al., 2012). Mollazade et al. (2012) have compared model accuracy of several machine learning methods for grading raisins based on visual images.

Among various machine learning methods, the Random Forests (RF: Breiman, 2001) algorithm has been regarded as one of the most precise prediction methods, having advantages such as ability to determine variable importance, ability to model complex interactions among predictor variables, and flexibility to perform several types of statistical data analysis including regression, classification and unsupervised learning (Cutler et al., 2007). The use of RF allows for a new way of modelling and extracting information from observation data, and thus contributes to a better understanding of a target system and mechanism that are, in general, complex and nonlinear. Its high predictive capability has been supported by previous comparative studies with other machine learning methods (Benito Garzón et al., 2006; Peters et al., 2007; Slabbinck et al., 2009; Kampichler et al., 2010; Pino-Mejías et al., 2010; Bisrat et al., 2012; Fukuda et al., 2013a). Also, it has been successfully implemented for yield estimation of agricultural products (Vincenzi et al., 2011; Fukuda et al., 2013b). However, to the best of our knowledge, no study has applied RF in food engineering.

This study aims to demonstrate the applicability of RF as a tool for estimating the internal quality of fresh mango fruit based on peel colour. The results are presented with a specific focus on the relationship between peel colour and fruit quality changes, considering the postharvest ripening process of mango fruit. Specifically, variable importance computed by RF was investigated to interpret the model results. As such, the method of how to interpret the model results can be illustrated (i.e. specification of the important colour parameters for modelling the quality of fresh mango fruit).

2. Materials and methods

Two commercial mango cultivars, having different fruit properties and grown in intensively managed orchards in Thailand and Japan, were used to test the applicability of proposed approach and compare the difference between the cultivars. In fruit quality monitoring, the changes in peel colour and fruit quality were observed before and after distribution under three postharvest conditions with respect to temperature, while distribution condition was the same across the postharvest conditions. Based on the observation data of changes of peel colour and fruit quality, a set of RF models were developed and evaluated with respect to model accuracy and variable importance computed by the RF models.

2.1. Fruit sample collection

Fresh fruit samples of two mango cultivars, namely ‘Nam Dokmai’ (Fig. 1a) and ‘Irwin’ (Fig. 1b), were collected from an intensively managed orchard in Phitsanulok, Thailand (16°33′ N, 100°37′ E, 64 m a.s.l.) and in Okinawa, Japan (26°10′ N, 127°41′ E, 18 m a.s.l.), respectively. These orchards were selected because they produce high quality mangoes for export to major markets in Japan. The cultivars used differed considerably with respect to phenotypic attributes. Each represented one of the two races in which mangoes are classified, namely subtropical (monoembryonic) and tropical (polyembryonic). Subtropical varieties usually exhibit the ‘bicolor’ characteristic of green peel with a red shoulder, while the tropical varieties are normally considered to be ‘all yellow’, although some mango varieties can even exhibit unordinary hues of red and purple. Irwin is described as subtropical, exhibiting development of a red shoulder and Nam Dokmai is included in the tropical yellow varieties (Litz and Gomez-Lim, 2005).

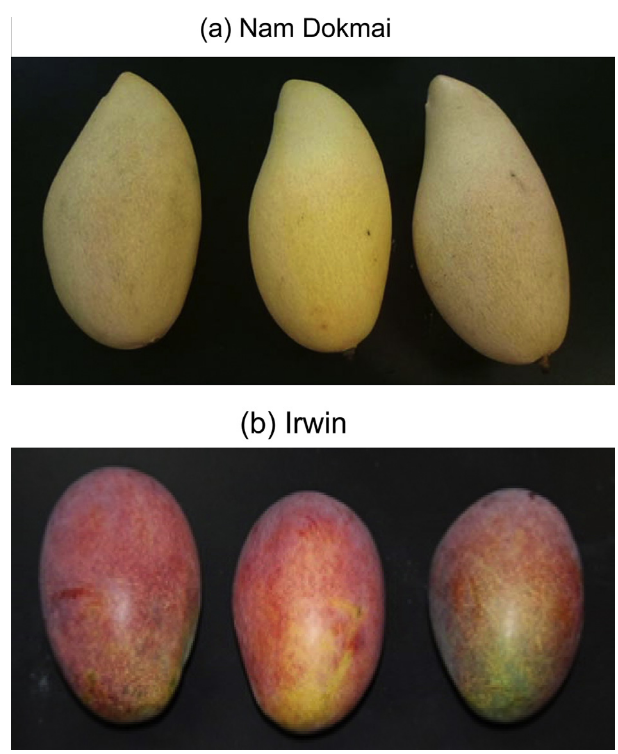

Fig. 1. Photographs showing the (a) Nam Dokmai and (b) Irwin samples after distribution.