Agricultural Water Management 116 (2013) 142–150

Contents lists available at SciVerse ScienceDirect

Agricultural Water Management

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

Random Forests modelling for the estimation of mango (*Mangifera indica* L. cv. Chok Anan) fruit yields under different irrigation regimes

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

[a] Institute of Tropical Agriculture, Kyushu University, Hakozaki 6-10-1, Fukuoka 812-8581, Japan
[b] Institute of Agricultural Engineering, Universität Hohenheim, 70593 Stuttgart, Germany
[c] 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
[d] Faculty of Agriculture, Kyushu University, Hakozaki 6-10-1, Fukuoka 812-8581, Japan
[e] School of Agro-Industry, Mae Fah Luang University, Chiang Rai 57100, Thailand

ARTICLE INFO

Article history:
Received 2 March 2012
Accepted 8 July 2012
Available online 28 July 2012

Keywords:
Data mining
Yield estimation
Water management
Knowledge extraction

ABSTRACT

'Chok Anan' mangoes, mainly produced in Northern Thailand, are appreciated for their light to bright yellow colour and sweet taste. Because fruit development of the on-season mangoes occurs during the dry season, farmers have to irrigate mango trees to ensure high yields and good quality. Therefore, it is important to understand the effects of water supply on the yield of mango fruit for better control and effective use of limited water resources. In this study, we aim to demonstrate the applicability of Random Forests (RF) for estimating mango fruit yields in response to water supply under different irrigation regimes. To cope with the variability of mango fruit yields observed in the field, a set of RF models was developed to estimate the minimum, mean and maximum values for each of the mango fruit yields, namely "total yield" and "number of marketable mango fruit". In RF modelling, a combination of 10-day rainfall and irrigation data was used as model input in order to evaluate the effects of water sources on the mango fruit yields. The RF models accurately estimated the maximum and mean values of mango fruit yields, and showed moderate accuracy for the minimum mango fruit yields. The variable importance measure computed in the RF calculation suggested that the timing of water supply affects the mango fruit yields whereby rainfall and irrigation have different effects on the mango fruit yields. This case study on the estimation of mango fruit yields demonstrates the applicability of RF in the field of agricultural engineering, with a specific focus on water management. The model performance and the information retrieved from the RF models allow for precise modelling and the development of improved management practices in target agricultural systems.
© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Thailand is one of the most important mango producers and exporters in the world. Out of the multitude of mango varieties grown worldwide, the following five varieties have major export potential in Thailand: 'Nam Dok Mai', 'Nam Dok Mai Si Thong', 'Maha Chanok', 'Chok Anan', and 'Khiao Sawoei' (Chomchalow and Songkhla, 2008; DOA, 2012). While Nam Dok Mai mangoes are mainly exported to Japan (Thailand's main export market for fresh mango), Chok Anan mangoes are mainly exported to Singapore, Hong Kong, and other Asian countries. One feature that makes Chok Anan a suitable cultivar for export is its thick skin, which helps to reduce losses during transport (Tanongjid et al., 2010). Furthermore, the moderate ripening potential of this variety makes it most suitable for long supply chains based on optimised storage and subsequent ripening (Kienzle et al., 2011). In terms of sensory properties and carotene content, Chok Anan was found to be highly suitable for processed mango products (Mahayothee et al., 2007).

Regarding production, Chok Anan is unique in its ability to flower all year round (COA, 2011), allowing for up to three staggered harvests. However, a large share of the harvest is produced during the dry season when high radiation promotes fruit growth while growth of fungi and pest infestation pressure are reduced. Even though mango is a drought-resistant crop, mango fruit of export quality can only be obtained using irrigation. The quality grades of export mangoes are defined by colour, shape, aroma and fruit size. For premium-grade mangoes, farmers can obtain 30–50% higher revenues (Chomchalow and Songkhla, 2008). Irrigation enhances fruit size and has a strong influence on the marketable yield. It is,

* Corresponding author. Present address: Hohenheim Office, 5th Floor, New Building, Faculty of Agriculture, Chiang Mai University, Chiang Mai 50200, Thailand. Tel.: +66 53 944647; fax: +66 53 893099.
E-mail addresses: Wolfram.Spreer@gmx.net, spreerwo@uni-hohenheim.de (W. Spreer).

0378-3774/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.agwat.2012.07.003

thus, important to understand the effects of the timing of water supply on the yield response of a target cultivar in order to develop a better production system and improved management practices for water resources.

Two modelling approaches can be used to evaluate the relationship between the water supply and yield response of a target cultivar: a process-based modelling approach and a statistical (correlative) modelling approach. In general, a process-based modelling approach aims to develop a model based on the ecophysiological processes in response to environmental stresses (e.g. drought and heat stress) and nutrient conditions (e.g. the timing and amount of fertilizer application). A statistical modelling approach, such as data mining, aims to develop a model based on the relationship between the yield response of a target cultivar and various environmental conditions observed in a target system. On the one hand, the process-based approach requires detailed knowledge of the ecophysiological processes of a target cultivar (see Léchaudel et al., 2005a, 2007 for mango), hence the accuracy of the model depends on the knowledge and data available for the modelling, and the uncertainties underlying them. On the other hand, the statistical modelling approach can be employed whenever observation data on the yield response and environmental conditions are available (see de Bie, 2004; Durán Zuazo et al., 2004; González et al., 2004; Silva et al., 2009; Spreer et al., 2009b for regression models for mango). The model accuracy depends on the relationship between the input and output variables and the quality of data used for the model development. In contrast to the process-based models consisting of highly complex processes, the statistical modelling approach is more useful for its simplicity and interpretability in extracting information from the observation data. The information retrieved from a statistical model can be tested using existing knowledge and then further incorporated into process-based models. Therefore, the use of statistical models can contribute to the accumulation of knowledge on the complex interactions between the yield response of a cultivar and environmental conditions, such as water supply.

Recently, the application of computational intelligence and machine learning methods has been increasingly recognised in agricultural engineering (see Huang et al., 2010 for a review). However, few studies have applied such computational methods to estimate a fruit yield. For instance, artificial neural networks (ANNs) have been applied to estimate citrus yield from airborne hyperspectral images (Ye et al., 2006), and to model greenhouse tomato yield from automated crop monitoring data (Ehret et al., 2011). Delgado et al. (2009) applied fuzzy association rules to extract information useful for linking management practices with the quantity and quality of olive fruit. For their high predictive capability, advanced computational methods such as machine learning have high potential for modelling nonlinear systems in agriculture. Among various machine learning methods, Random Forests (RF; Breiman, 2001) has been widely applied to many disciplines because of the following advantages: (1) very high accuracy; (2) a novel method of determining variable importance; (3) ability to model complex interactions among predictor variables; (4) flexibility to perform several types of statistical data analyses including regression, classification, survival analysis and unsupervised learning; and (5) an algorithm for imputing missing values (Cutler et al., 2007). The highly predictive capability of RF has been supported by previous comparative studies with other machine learning techniques (Benito Garzón et al., 2006; Slabbinck et al., 2009; Kampichler et al., 2010; Pino-Mejías et al., 2010; Bisrat et al., 2012). However, so far no study has applied RF to estimate the yield of agricultural products (but see Vincenzi et al., 2011 for an RF application in clam fisheries).

The present study aims to demonstrate the applicability of RF for estimating the yield of mango fruit (*Mangifera indica* L. cv. Chok Anan) produced using four different irrigation treatments. A set of RF models was developed to estimate the minimum, mean and maximum values for each of fruit yield parameters, namely the total yield and the number of marketable mango fruit, based on 10-day rainfall and irrigation data. Moreover, the variable importance measure computed in the RF model was used to evaluate the effects of the timing of water supply (i.e. rainfall and irrigation) on mango fruit yields. Finally, we discuss the applicability of RF based on accuracy and the information retrieved from the present models.

2. Materials and methods

2.1. Irrigation treatment

Irrigation experiments were performed from 2005 to 2007 by using 196 fourteen-year-old 'Chok Anan' mango trees grafted on 'Talap Nak' rootstocks and spaced 4 m × 4 m in the orchard of Mae Jo University, Chiang Mai, Thailand (18.53°N, 100.03°E, 350 m ASL). A detailed description on the experimental set up (e.g. irrigation scheduling) as well as environmental conditions (e.g. soil property and weather conditions) was published by Spreer et al. (2007, 2009b).

Water requirements for irrigation were calculated as potential crop evapotranspiration (ET_c) according to the modified Penman–Monteith equation (Allen et al., 1998) with a crop coefficient (k_c) value of 0.8, for which climatic data from the meteorological station of Mae Jo University were used.

The trees were organized in a randomized block design comprising four repetitive blocks subdivided into eight fields. Mango fruit yields, namely total yield and the number of marketable mango fruit, were evaluated under four irrigation treatments: (a) full irrigation as the control (FI = 100% of ET_c), (b) regulated deficit irrigation (RDI = 50% of ET_c), (c) partial rootzone drying (PRD = 50% of ET_c) applied to alternating sides of the root system, and (d) no irrigation (NI).

In the PRD treatment in 2006 and 2007, the timing of the alternation of the irrigated and dried sides was set every 2 weeks, according to the method by Stoll et al. (2000). In 2005, the irrigated sides should have been changed at the occurrence of drought stress. However, stomatal aperture, a drought stress indicator, did not decrease due to early rainfall which interfered in the drying cycle of the PRD treatment. Thus, the alternation occurred only once at 2 weeks prior to harvest.

In the subsequent analysis, the 10-day rainfall and irrigation data were used for the estimation of mango fruit yields. The observations were recorded for 110 days, resulting in eleven sets of 10-day rainfall and irrigation data for each test year (Fig. 1). The use of 10-day data allows for assessing the effects of the timing of water supply under the different irrigation treatments.

2.2. Yield measurements

At harvest, the mango fruit yields, namely total yield and the number of marketable mango fruit, were determined separately for each tree using a mechanical balance. The harvested fruit was counted and classified by weight according to the recommendation of the Ministry of Agriculture of Thailand (Table 1). The classification scheme as well as the results on mango fruit yields and size distribution for the years 2005–2007 has been published by Spreer et al. (2007, 2009b). Mango fruit classified in the weight classes extra, 0, and 1 (weighing over 300 g) was defined as marketable mango fruit for export. To cope with the variability of mango fruit yields observed within and between irrigation treatments and between years (Fig. 2; see also Spreer et al., 2007, 2009b

Download English Version:

https://daneshyari.com/en/article/4478974

Download Persian Version:

https://daneshyari.com/article/4478974

Daneshyari.com