Random Forests as a Tool for Analyzing Partial Drought Stress Based on CO₂ Concentrations in the Rootzone of Longan Trees

Shinji FUKUDA¹, Wolfram SPREER², Winai WIRIYA-ALONGKORN³, Klaus SPOHRER⁵, Eriko YASUNAGA⁴ and Chantalak TIYAYON⁵

¹ Institute of Agriculture, Tokyo University of Agriculture and Technology, 3−5−8 Saiwai-cho, Fuchu, Tokyo 183−8509, Japan
² Institute of Agricultural Engineering, University of Hohenheim, Garbenstrasse 9, 70599 Stuttgart, Germany
³ Faculty of Agriculture, Chiang Mai University, 239 Huay Kaew Road, Mueang District, Chiang Mai 50200, Thailand
⁴ Department of Soil Resources and Environment, Mae Jo University, Chiang Mai, Thailand
⁵ 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

(Received October 22, 2017; Accepted December 24, 2017)

INTRODUCTION

In the Northern part of Thailand longan (Dimocarpus longan Lour.) is one of the major fruit crops. As a substantial part of longan fruit is conserved by drying, most of the longan production takes place during the dry season, the so called “on-season”. This means that most producers use artificial flower induction by KClO₃ for homogenizing the harvesting time, rather than for “off-season” production. This makes them more independent from the weather and alternate bearing. However, there is a series of factors, which affect successful flower induction, such as temperature, radiation or previous applications (Ongprasert et al., 2010). In recent years, farmers have encountered a series of very cold winters with nighttime temperatures below 10°C, which have affected the flowering negatively.

As longan trees are particularly sensitive to drought during the flowering and early fruit development stages, in growing regions which have a distinct summer rainfall pattern, high fruit yields can only be obtained using irrigation (Ongprasert et al., 2007). This is the case in Thailand, where under the impression of a changing climate the year-to-year variability of rainfalls increase dramatically, imposing new challenges to fruit production (Schulze et al., 2013). In the long run, limited water resources are an obstacle to increase longan production. It is crucial that excessive irrigation is avoided in order to use the existing water resources at an optimal level. At the same time good yields in terms of quantity and fruit quality can only be obtained when water deficit is avoided, as longan reacts sensitively to drought by fruit yield reduction and fruit cracking. In order to guarantee sufficient water-supply, state-of-the-art irrigation planning relies mainly on monitoring of soil moisture and/or climatic water balance. The reason is that soil and climate data can be continuously monitored, enabling the irrigator to have a constant overview on the performance of the irrigation system. For instance, Yasunaga et al. (2016) demonstrated a semi-real-time monitoring system for field experiments of precise irrigation in mango production. The monitoring of actual crop related data — such as stomatal conductance, leaf water potential or chlorophyll fluorescence — is mainly used in science, rather than actual irrigation management. This is due to the nature of the measuring principles, which are intermittent, time consuming, strongly invasive or even destructive. However, it is hypothesized that for a good irrigation management, besides soil moisture and climate data, monitoring of plant responses is important.

The challenge is therefore, to identify methods of
monitoring plant responses in a non- or little invasive way with the potential for continuous observations. So far, monitoring of canopy temperature by thermal imaging was found to have a high potential, as data are obtained by means of remote sensing. Canopy temperature is correlated to stomatal aperture. It was shown that this technique has a potential to be applied in drought stress monitoring of longan trees (Wiriyalongkorn et al., 2013). However, continuous monitoring of canopy temperature is difficult in practice. Fixed installed thermal cameras can be used in greenhouses (Grant et al., 2006; 2007), while the use of unmanned aerial vehicles (UAV) may be more suitable for on-site monitoring. For instance, Watanabe et al. (2017) employed a UAV as an improved tool for precise monitoring of sorghum plant height and predictive modelling for genomics-assisted breeding. Both monitoring and data evaluation in these cases are demanding in terms of equipment and automation software. Moreover, changes in stomatal aperture or photosynthesis happen at a rather advanced stage of plant stress and are, therefore, more useful for the retrospective description of drought stress rather than reliable indicators for stress at an early stage.

We analyzed the monitoring of root-born CO₂ for its potential to estimate the water status of the plant. If applicable, the CO₂ monitoring in the rootzone bears the potential to be a parameter which can be detected in a similar way as soil moisture with a minimum level of invasion and with a good potential for automation (Spoehr et al., 2015). Root respiration is an important indicator for growth and health of a plant as all vital functions, such as root growth, water and nutrient uptake, depend on energy from root respiration. As respiration is difficult to determine in an agricultural field, CO₂ efflux is measured at the surface or at low depth. This procedure is commonly accepted as it was shown that the superficial CO₂ efflux is correlated with root respiration under normal conditions. Irrigation, however, changes the conditions in the rootzone: Where all pores are full of water, a layer with saturation like conditions is formed, which exerts high resistance to gas exchange. This is known as the wetting front, which moves downward during the irrigation event. Under these special conditions it is no longer possible to strictly correlate CO₂ efflux to root respiration. This was documented by Bouma and Bryla (2000) who showed that by determination of the CO₂ efflux, root and soil respiration were underestimated after an irrigation event. The effect was observed to be longer lasting at a finer texture. This may be either due to the fact that the wetting front moves more slowly, or because of the expansion of the wetting front, which is higher at a finer texture.

With respect to the interaction between water availability and root respiration, correlations have been shown in several studies, mainly based on the effect that drought stress reduces plant growth, which ultimately affects root growth. Soil water content influences availability of carbohydrates in the soil and thereby root borne CO₂ production (Liu and Li, 2006). Reactions of roots on water stress are closely related to adaptation mechanisms on changing temperature regimes. A study showed a high temperature dependency of CO₂ efflux from the rootzone of citrus trees at good water supply, while there was no influence of temperature on water-stressed trees (Bryla et al., 1997). In grape vine it was observed that at a higher temperature root respiration in the drying soil decreases faster and more than at a lower temperature, which means that roots are more sensitive to drought if exposed to heat stress (Huang et al., 2005). At the often occurring combination of water and heat stress the trees have little chance to adapt to the changing conditions. There is a direct correlation between CO₂ concentration in the rootzone and water supply by irrigation as shown by Spoehr et al. (2015) who monitored CO₂ concentration at a depth of 20 cm in the rootzone of young apple trees. In contrast, CO₂ monitoring for our previous study took place in a depth profile of 5–30 cm (Wiriyalongkorn et al., 2016). It turned out that the closer to the surface, there is a higher influence of air temperature on the measurement. The correlation of water supply and CO₂ concentration suggested earlier detectable changes in rootzone CO₂ after changes in water supply as compared to stomatal conductance.

Monitoring the CO₂ dynamics in the rootzone and correlating them to water supply under changing environmental conditions, especially temperature, requires the handling of large data-sets and complex analyses of the interactions. Due to the complex eco-physiological processes, detailed theoretical models cannot be developed with commonly observed data such as weather and soil moisture conditions. The main problem with a standard statistical approach in our previous study was found when analyzing the changes over time that were highly variable in their frequency and magnitude. A direct comparison between treatment and control made those difference visible but with a low coefficient of determination. Data-driven models, specifically machine learning, have been widely used for analyzing input-output relationships in various research fields. The application of machine learning in agricultural sciences as such is not so new. However, advances in information and communication technology (ICT) enable collecting, storing and communicating a huge amount of data from field and laboratory, thereby fostering predictive modelling and knowledge extraction based on observation data. Random forests (RF; Breiman, 2001) is one of the most commonly used methods for its predictive performance and useful functions for knowledge extraction. This paper demonstrates how RF-based models can relate field observed data with CO₂ concentrations in the rootzone. Two kinds of information, namely variable importance and response curves, are used for better understanding the complex interaction between CO₂ concentrations and water supply.

MATERIALS AND METHODS

Experiments and data acquisition
The experiment was carried out at Mae Jo University, Chiang Mai Province, Thailand (18°53′40″N, 99°00′58″E) from November 1st 2011 to April 30th 2012 (Wiriyalongkorn et al., 2016). Five-year-old longan trees with a

S. FUKUDA ET AL.
split root system grown in two separate cement containers (diameter 0.8 m, height 0.5 m) were kept under a transparent plastic shelter with free air circulation. Washed sand was used as growing media. The tree height was approx. 2.0 m and the average canopy diameter was 1.8 m. The daily water consumption was determined as 10 l/tree. For the duration of the experiment irrigation was applied daily; 5 trees, which served as control (CO), received 5 l of irrigation water in each of the two pots. Another five trees received 5 l of irrigation water in one pot only, either to the left side of the rootzone (L) or its right side (R). The irrigated and the dry side were changed in a two-weeks-interval.

During the experiment air temperature (T), relative humidity (RH), wind speed and atmospheric pressure were monitored by a portable weather station (PCE-FWS 20, PCE Group, Germany). The vapor pressure deficit (VPD) was calculated as the difference between the saturated vapour pressure at temperature T (eSAT(T)) and the actual vapour pressure (e) using the following equations:

\[ e(T) = 0.6108 \times \exp\left(\frac{17.27 \times T}{T+237.3}\right) \]

\[ e = e(T) \times \frac{RH}{100} \]

Soil moisture was determined by time domain reflectometry (TDR) using a 1502C cable tester (Tektronix, USA) one hour before and after each irrigation event. The CO2-concentration in the rootzone measured by the use of a TES 1370 NDIR CO2 Meter (TES Electronic Corp., TW), which were placed in a sealed container connected to a 30 cm long 50-mm PVC pipe with a perforated area of (94 × 20 mm²) introduced into the rootzone. In the container there was no gas exchange other than diffusion from the rootzone. The increase in CO2 concentration as compared to the initial (ambient) concentration in the container was measured after 20 min. A detailed description of the experimental set-up was described by Wiriya-Alongkorn et al. (2016).

**Random forests modelling**

We applied RF for modelling CO2 concentration in the rootzone of longan trees, in which minimum, mean and maximum values of ambient temperature (°C) and relative humidity (%) as well as vapor pressure deficit (kPa) and soil moisture content (m³/m³) were used as model input. Model performance was assessed based on root mean squared error (RMSE), Nash-Sutcliffe efficiency (NSE) and Pearson’s correlation coefficients (COR):

\[
\text{COR} = \frac{\sum \left( Y_o \times Y_m - \sum \left( \frac{Y_o \times Y_m}{\sum Y_o \times \sum Y_m} \right) \right)^2}{\sqrt{\sum Y_o^2 \times \sum Y_m^2}} \times \frac{1}{\sum Y_o^2 \times \sum Y_m^2} \quad (1)
\]

\[
\text{NSE} = 1 - \frac{\sum \left( Y_o - Y_m \right)^2}{\sum \left( Y_o - \overline{Y_o} \right)^2} \quad (2)
\]

\[
\text{RMSE} = \sqrt{\frac{\sum \left( Y_o - Y_m \right)^2}{\sum \left( Y_o - \overline{Y_o} \right)^2}} \quad (3)
\]

where \( Y_o \) and \( Y_m \) are the observed and modelled CO2 concentration of the data point \( i \) (\( i = 1, 2, \cdots, N \)), \( \overline{Y_m} \) is the mean observed CO2 concentration, and \( N \) is the size of the data set. Specifically, the NSE takes a value in the range \([-\infty, 1]\), of which an NSE value of 1 means a perfect fit. Of these measures, the mean and standard deviation from 50 different sets of initial conditions in RF computations were used as the measure of the accuracy and the variability of model structures, respectively. We used the randomForest package (Liaw and Wiener, 2002) of the R software (R Development Core Team, 2011), in which the default settings were applied. After removing missing observation data, we obtained 589 data for modelling CO2 concentrations in the rootzone.

In this study, variable importance and response curves derived from RF computations were used as a tool of knowledge extraction (Fukuda et al., 2013). We employed variable importance, computed as the percent increment of mean squared error (MSE) using the `varimp` function, for understanding which variables are of importance when modelling CO2 concentration in the rootzone. Response curves were derived for each input variable using the ‘parPlot’ function of RF for understanding which conditions control CO2 concentration in the rootzone.

**RESULTS**

**Soil moisture and irrigation regime**

While soil moisture content showed distinct differences between the partial root-zone drying (PRD) and control treatment (Fig. 1), no apparent difference was observed in CO2 concentrations (Fig. 2). Due to the treatment conditions, soil moisture content in the PRD treatment was different between two sides (i.e. left and right) but CO2 concentration was different between trees, masking differences between the sides to a large extent. As a rough trend in the control treatment, soil moisture content increased toward January and gently decreased afterwards, while CO2 concentrations remained low between December and February. Detailed discussion on the observation data can be found in Wiriya-Alongkorn et al. (2016).

**Random forests**

The RF models showed moderate performance on modelling CO2 concentrations in the rootzone of longan trees under different irrigation regimes (Fig. 3): RMSE = 301.0 ± 0.320, NSE = 0.455 ± 0.001; COR = 0.680 ± 0.001 (mean ± standard deviation). The variability in the model structure was small as shown in the standard deviation in model performance. The RF model overestimated the CO2 concentrations in case of lower CO2 concentrations while it underestimated CO2 concentrations in case of higher concentrations. In other words, it was difficult for the RF models to distinguish high and low CO2 concentrations in the rootzone from given input data. For instance, on a modelled CO2 concentration of 1,000 ppm, we can observe varying observed CO2 concentrations from 0 to 2,300 ppm (Fig. 3). This result partially indicates that insufficient information was available for the RF model to estimate CO2 concentrations in the rootzone.

Soil moisture content was found to be the most
important parameter when modelling CO₂ concentration in the rootzone (Fig. 4). The gaps in variable importance between soil moisture content and other variables were distinct but differences between other variables such as temperature-based and humidity-based variables were small. Variability in variable importance was relatively large in soil moisture content.

Response curves demonstrate how an input variable affects CO₂ concentrations in the rootzone of a longan tree (Fig. 5). Response in CO₂ concentrations correlated with maximum air temperature up to 38 °C and flattened out (Fig. 5a), while that for minimum air temperature remained below 800 ppm and showed a distinct increase around 18°C (Fig. 5c). In contrast, mean air temperature showed an increasing trend with some sudden changes around 24 and 26°C (Fig. 5b). Whereas maximum and minimum relative humidity showed slightly increasing trends, mean relative
humidity showed a decreasing trend (Fig. 5d-f). Daily temperature range exhibited a decreasing trend (Fig. 5g) and vapour pressure deficit showed a nearly flat response (Fig. 5h). The response in soil moisture content was the most distinct compared to the other variables (Fig. 5i), which exhibited a slight decrease in the lowest soil moisture content and increased up to around 0.12 where the curves were flattened out.

**DISCUSSION**

If monitoring the presence of CO2 in the rootzone has a potential to analyze plant water status, it is a novel method, which combines the advantages of soil moisture monitoring, such as non-destructive measurements and the option of continuous data acquisition. However, it is clear that not all the CO2 in the rootzone originates from the crop to be monitored and the production of CO2 in the rootzone itself is exposed to a variety of influencing factors. The experimental data showed that there is a significant correlation between the CO2 concentration in the rootzone and the soil moisture, but the simple coefficient of determination was low (Wiriya-Alongkorn et al., 2016). By a more complex analysis with random forests (RF), the correlation between CO2 concentration in the rootzone and soil moisture was further confirmed, showing that it is by far the most meaningful correlation as compared to the other factors monitored. Interestingly, the response curves showed the most pronounced response at soil moisture contents between 5% and 10% v/v, which corresponds to pF values between 4.2 and 2.8 in the analyzed substrate. The closer the value are to field capacity, the less response was found in the RF analyses. This result points to a threshold of physiological drought stress, which is expressed by root activity, so that beyond this threshold the CO2 production is at maximum level and no major changes are to be expected.

At the same time, it is also possible that the presence of water in the soil itself limits the gas-exchange to a certain degree, which implies that depending on the soil type, texture and layering there is a physical maximum of CO2 concentration in the rootzone, as well as a maximum level of gas exchange. In a dry soil the complete pore space is filled with gas. Thus, a change in the gas composition affects the surrounding soil body in a very short time. In the presence of water, the pore space is limited and the gas exchange is slowed down, especially when observing the situation under irrigation, where the advancing wetting front is a layer of saturation-like conditions that can stop the gas exchange completely. It is therefore important to determine CO2 concentration in the soil, without the direct influence of an irrigation event. Continuous measurements as done by Spohrer et al. (2015) indicated that the time of
variation of CO₂ concentration in the rootzone as an effect of changing soil moisture was shorter at lower soil moisture. Irrigation on a dry soil provokes an immediate change in CO₂ concentration, while changes at high water content are slower. That implies that irrigation changes soil moisture content at a faster rate than water extraction by the plant. Consequently, a wetting front is created which produces local peaks in CO₂ concentration in the root-zone. In this context the question about the soil depth, which is appropriate to measure, must be investigated.

In the experiment providing the data for this study, the main influencing factor on CO₂ concentration in the rootzone other than soil moisture was temperature. While a simple consideration of CO₂ data and temperature did not render a clear image, analyses using RF importantly pointed to the influence of low temperatures on the physiological responses of longan. This adds a new aspect to the monitoring of plant reactions in the agricultural practice, as in terms of climatic conditions, longan production in Northern Thailand is affected by low temperatures during the dry season. Cold stress plays an important role in natural flower induction of Thai longan varieties. In this analysis it was shown, that there is a pronounced response of CO₂ concentration at a minimum temperature above 18°C. This points out that this temperature level can be considered the threshold, where physiological processes in longan are affected. At the same time maximum temperatures have a much smaller influence on longan physiology and up to 35°C there is even an increase in response of CO₂ concentration to increasing temperatures. On the other hand, there is no indication that CO₂ concentration in the rootzone changes under heat stress, as even at maximum temperatures well above 45°C there is no change in the response.

In contrast to the temperature and soil moisture content, relative humidity of the air and, consequently, vapour pressure deficit (VPD) show almost no relationship with CO₂ concentration in the rootzone. This is in accordance with the observation that there is no correlation between the photosynthetic rate and the CO₂ concentration in the rootzone (Wiriya-Alongkorn et al., 2016). This may partly because, gas exchange is not yet affected under low levels of stress, but can only be detected at a more severe level of stress. The fact that a low level of stress can be assumed in the present study is reflected in soil moisture content, which decreased towards the end of the experiment.

More research is required in order to correlate changes in CO₂ concentration in the root zone to the beginning of drought stress. Stress levels below the level of severity that provoke stomatal closure often lead to a reduced biomass formation. However, so far it is not possible to determine this level by continuous plant-based measurements. Therefore, soil-moisture monitoring is used instead, which does not consider the individual plant performance. Rootzone CO₂ monitoring in contrast may have a potential to allow the determination of low levels of plant stress in real time and, therefore, may play an important role in deficit irrigation practices, such as partial rootzone drying, where reactions on light drought stress are exploited to improve water use efficiency (WUE).

CONCLUSIONS

In this paper, we applied random forests (RF) to estimate CO₂ concentrations in the rootzone of longan trees based on climatic and soil conditions. The results underlined the potential of RF as a tool of predictive modelling and knowledge extraction. Variable importance identified soil moisture content as the major factor influencing CO₂ concentrations in the rootzone, even when partly masked by changes in temperature. Response curves indicated that physiological changes of longan can be at daily minimum temperature around 18°C. Considering the eco-physiological responses of a plant, model performance can be improved by using more closely related parameters such as soil temperature. High contribution of soil moisture content to the CO₂ concentration found from RF modelling points towards the potential of a novel monitoring tool of plant water status. Continuous and precise measurement of plant status is an important step for the development of precise and adaptive management of a production system.

ACKNOWLEDGMENTS

This work was supported in part by the Grant-in-Aid for Scientific Research (B) (Grant No. 16H05800). Data acquisition was carried out in the framework of the collaborative research project SFB 564 “The Uplands Program” funded by the German Research Foundation (DFG).

REFERENCES

Bouma, T. J., Bryla, D. R. 2000. On the assessment of root and soil respiration for soils of different textures: interactions with soil moisture contents and soil CO₂ concentrations. Plant Soil 227: 215–221.

Breiman, L. 2001. Random forests. Machine Learning 45: 5–32.

Bryla, D. R., Bouma, T. J., Eissenstat, D. M. 1997. Root respiration in citrus acclimates to temperature and slows during drought. Plant Cell Environ. 20: 1411–1420.

Fukuda, S., De Baets, B., Waegeman, W., Verwaeren, J., Mouton, A. M. 2013. Habitat prediction and knowledge extraction for spawning European grayling (Thymallus thymallus L.) using a broad range of species distribution models. Environ. Modell. Softw. 47: 1–6.

Grant, O. M., Chaves, M. M., Jones, H. G. 2006. Optimizing thermal imaging as a technique for detecting stomatal closure induced by drought stress under greenhouse conditions. Physiol. Plant. 127: 507–518.

Grant, O. M., Tronina, L., Jones, H. G., Chaves, M. M. 2007. Exploring thermal imaging variables for the detection of stress responses in grapevine under different irrigation regimes. J. Exp. Bot. 58: 815–825.

Huang, X., Lakso, A. N., Eissenstat, D. M. 2005. Interactive effects of soil temperature and moisture on Concord grape root respiration. J. Exp. Bot. 56: 2651–2660.

Liaw, A., Wiener, M. 2002. Classification and regression by random forest. R News 2: 18–22.

Liu, H. S., Li, F. M. 2006. Effects of shoot excision on in situ soil and root respiration of wheat and soybean under drought.
stress. Plant Growth Regul. 50: 1–9.

Ongprasert, S., Spreer, W., Wiriyalongkorn, W., Ussahatanonta, S., Köller, K. 2007. Alternative techniques for water-saving irrigation and optimised fertigation in fruit production in Northern Thailand. In “Sustainable Land Use in Mountainous Regions of Southeast Asia: Meeting the Challenges of Ecological, Socio-Economic and Cultural Diversity” (ed. by Heidhues, F. J., Herrmann, L., Neef, A., Neidhart, S., Pape, J., Sruamsiri, P., Thu, D. C., Vallé Zarate, A.), Springer. Berlin/Heidelberg/New York/London/Paris/Tokyo, p 120–133.

Ongprasert, S., Wiriyalongkorn, W., Spreer, W. 2010. The factors affecting longan flower induction by chlorate. Acta Hortic. 863: 375–380.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Schulze, K., Spreer, W., Keil, A., Ongprasert, S., Müller, J. 2013. Mango (Mangifera indica L. cv. Nam Dokmai) production in Northern Thailand - Costs and returns under extreme weather conditions and different irrigation treatments. Agric. Water Manage. 126: 46–55.

Spohrer, K., Zia-Khan, S., Spreer, W., Müller, J. 2015. Real-time detection of root zone-CO2, and its potential for irrigation scheduling. Landtechnik — Agric. Eng. 70: 150–156.

Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., Yano, K., Tokunaga, T., Fujiwara, T., Tsutsumi, N., Iwata, H. 2017. High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling. Front. Plant Sci. 8: 421.

Wiriyalongkorn, W., Spreer, W., Ongprasert, S., Spohrer, K., Pankasemsuk, T., Müller, J. 2013. Detecting drought stress in longan tree using thermal imaging. Maejo Int. J. Sci. Technol. 7: 166–180.

Wiriyalongkorn, W., Spreer, W., Ongprasert, S., Spohrer, K., Müller, J. 2016. Influence of water supply on CO2 concentration in the rootzone of split-root potted longan trees. Lect. Notes Comput. Sci. 9835: 346–356.

Yasunaga, E., Fukuda, S., Spreer, W., Takata, D. 2016. Online monitoring system on controlled irrigation experiment for export quality mango in Thailand. Lect. Notes Comput. Sci. 9835: 328–334.