Analysis of Land Characteristics for Efficient Irrigation Development of Sugarcane Growing Areas in Khon Kaen Province, Thailand

Preecha Kapetch¹, Kazuhiro Sakai², Kobkiat Pisanjaroen³, Taksina Sansayawichai⁴, Hiroyuki Cho⁵, Shinya Nakamura⁶ and Tamotsu Nakandakari⁷

Abstract: North-eastern Thailand has little rainfall and requires efficient irrigation development to enhance stable sugarcane production. However, identifying the highest priority areas for irrigation development is complex because the benefit derived from irrigation development depends on rainfall, available irrigation water, and soil characteristics. We used the CANEGRO model to simulate the sugarcane yield of existing cultivation areas under rainfed and irrigated conditions, taking into account actual weather and soil type. We then calculated the benefit of the irrigation development using the simulation results and actual data for groundwater well capacities, sugarcane prices, and irrigation development and running costs. We then analysed the results of the benefit calculation by ABC analysis and the decision tree method. The decision tree analysis confirmed that well capacity most influenced benefit. Areas with higher rainfall had high yields under rainfed condition, so the benefit from irrigation was small (or even negative as the cost of irrigation exceeded the increased income). A notable finding was that low soil available water content resulted in low yields in both rainfed and irrigated conditions, and high available water content resulted in high yields under rainfed conditions; therefore, both low and high available water content resulted in low benefit from irrigation development.

Keywords: CANEGRO; Sugarcane; ABC analysis; Decision tree; Irrigation development

1 Introduction
Thailand is the fourth ranked sugarcane producer and the second ranked sugar exporter in the world. The sugarcane growing area in Thailand is 1.7 million ha, and 117 million t of sugarcane is produced per year (OCSB 2015). Production is also steadily increasing because of the increasing area under sugarcane cultivation. However, the productivity of land under cultivation is not improving. Because 80% of sugarcane fields are rainfed, some areas face drought stress. Moreover, the sugarcane yield is unstable because of rainfall variability. Irrigation development is therefore an important objective to improve sugarcane production level and stability. However, irrigation development requires public investment, which demands efficient use of the funds. It is therefore important to prioritize areas for development by taking into account variations in climate and soil. Crop models are a powerful tool that could be employed to address this issue of spatial prioritization for investment. The CANEGRO model is widely used to simulate sugarcane production and has been continually improved since the 1970s (Inman-Bamber et al. 2002; Jintrawet et al. 1997; O’Leary 2000; Singels and Bezuidenhout 2002; Singels et al. 2005). CANEGRO has been combined with DSSAT ver. 4.5 (Singels and Bezuidenhout 2002) to simulate the effect of water stress, but not nitrogen stress. Preecha et al. (2015) calibrated CANEGRO for sugarcane production in north-eastern Thailand and found that it could validly simulate the aboveground dry matter and stalk dry matter of three cultivars of sugarcane under irrigated and rainfed conditions.

In this study, we used CANEGRO to simulate sugarcane yield in current sugarcane production areas, first under rainfed conditions and then assuming automated irrigation. We then calculated the benefit that would be obtained from irrigation development using the results of the simulation and well capacity data (irrigation water availability). We sorted the results by the benefit per area and calculated the total benefit. We then used ABC analysis of the total benefit to highlight the priority areas. Finally, we conducted decision tree analysis to more clearly characterize the conditions that provide the best return from irrigation development.

2 Materials and methods
2.1 Spatial data and land units for simulating sugarcane production
The study employed spatially distributed digital data of current sugarcane growing areas, soil groups, weather stations, and well capacities in Khon Kaen province. The data were collated and manipulated in the ArcView 3.2a (Esri,
Redlands, CA, USA) geographic information system (GIS) software. The digital map (polygons) of sugarcane growing area (Figure 1) was obtained from the Office of the Cane and Sugar Board (OCSB 2013, 2015). The data were collected in the year 2012/13. Soil data were obtained from the Land Development Department, Ministry of Agriculture and Cooperatives (LDD 2000). The attributes of the soil data were soil group and land type arranged in polygons. The sugarcane growing area included 28 soil groups (Figure 2). The majority of the soils in the sugarcane growing area were sandy loam. The soil data were used as input data to the CANEGRO model. The available water content, saturated hydraulic conductivity, and bulk density of each soil group are shown in Table 1. The weather station data were obtained from the Department of Meteorology (TMD 2010). There are seven weather stations in Khon Kaen province (WS-1 to WS-7; Figure 3). The delineation of areas allocated to each station was determined by the Thiessen method. The 30-year-average annual rainfall of each weather station is shown in Table 2.

We defined the simulation conditions of the sugarcane growing area according to the soil group and weather station. Overlaying the soil map of Figure 2 on the weather station map of Figure 3 produced 82 simulation conditions. Each land condition was considered discrete for sugarcane production because of its unique combination of soil group and weather conditions. We simulated the sugarcane production for each of these 82 simulation conditions. The groundwater map for well capacity was obtained from the Department of Groundwater Resources (DGR 2014), and groundwater yield (m³ hr⁻¹) in the sugarcane growing area is shown in Figure 4. Four ranges of the groundwater yield shown in the original map of DGR were <2 m³ hr⁻¹, from 2 to 10 m³ hr⁻¹, from 10 to 20 m³ hr⁻¹, and...
Therefore, we defined four categories of well capacity. WC-1, -6, -15, -25 represent well capacity of 1, 6, 15, and 25 m³ hr⁻¹. Overlaying the groundwater yield map of Figure 4 on simulation conditions produced 179 land units.

2.2 Model simulations

We applied the CANEGRO model bundled with DSSAT ver. 4.5 (Hoogenboom et al. 2011) to simulate sugarcane yields for 30 years from 1980 to 2009. The input data consisted of crop management factors, crop genetic coefficients, daily weather data (rainfall, maximum and minimum temperature, and solar radiation), and physical and chemical properties of the soils. We used the input data for crop management recommended by the Department of Agriculture. The planting density was 5 plants m⁻². Sugarcane was planted on 15 October and harvested on 15 December of the next year. The genetic coefficients of the sugarcane variety KK3 were used, because this is a popular variety in Khon Kaen.

The conditions used for the simulations were the rainfed condition (RFC) and the automatic irrigation condition (AUIC) set in DSSAT. In the AUIC, irrigation water was supplied to 100% of the available water when the water content decreased to 50% of the available water. Adequate nitrogen fertilizer was provided in both conditions to remove soil fertility as a variable.

We simulated the sugarcane production of the 82 simulation conditions for the period 1980 to 2009. The output data extracted for the subsequent analyses were the sugarcane yield and the amount of irrigation water consumed.

2.3 Benefits calculation

We calculated the yield gap, the yield under AUIC minus that under RFC, for each simulation condition and overlaid it with the well capacity. For each discrete area (polygon), the benefit values were calculated from

\[ BF = IC - CS \]

\[ IC = GP \times A \times P_s \]

\[ CS = FC + VC \]

\[ FC = K_{FC} \times A \]

\[ VC = K_{VC} \times T_i \times n \]

\[ T_i = \frac{10 \times IR \times A}{W_C \times 10,000} \]

\[ n = \frac{32,000}{20} \]

where BF is the benefit in Thai baht (THB), IC is the income (THB), CS is the cost (THB), GP is the yield gap (kg ha⁻¹), A is the area (ha), \( P_s \) is the price (THB kg⁻¹), FC is the fixed cost (THB ha⁻¹), VC is the variable cost (THB hr⁻¹), \( K_{FC} \) (THB ha⁻¹) and \( K_{VC} \) (THB hr⁻¹) are coefficient, \( T_i \) is the hours of pump operation, \( n \) is the number of pumps installed in the one polygon, \( IR \) is the simulated irrigation water (mm), and \( W_C \) is the well capacity (m³ hr⁻¹). The fixed cost is the initial investment cost of developing the irrigation scheme. The variable cost is the running cost of providing irrigation water. In this study we assumed that the same pump system installed every 32,000 m² was used. We used 7,603 THB ha⁻¹ for \( K_{FC} \). That price included costs of the well construction, the pump, and the drip line (OCSB and KKU 2015). Then, we used 5 THB hr⁻¹ for \( K_{VC} \) (OCSB and KKU 2015).

| SG  | AWC (mm mm⁻¹) | Ks (cm hr⁻¹) | BD (g cm⁻³) |
|-----|---------------|--------------|-------------|
| 1467| 0.13          | 0.06         | 1.44        |
| 1468| 0.13          | 0.34         | 1.44        |
| 1469| 0.14          | 0.15         | 1.43        |
| 1470| 0.14          | 3.01         | 1.49        |
| 1471| 0.13          | 0.61         | 1.45        |
| 1472| 0.12          | 0.40         | 1.61        |
| 1473| 0.12          | 1.54         | 1.41        |
| 1474| 0.12          | 0.40         | 1.58        |
| 1475| 0.12          | 0.40         | 1.40        |
| 1476| 0.12          | 0.87         | 1.55        |
| 1477| 0.12          | 6.10         | 1.40        |
| 1478| 0.12          | 1.32         | 1.43        |
| 1479| 0.12          | 0.28         | 1.41        |
| 1480| 0.12          | 0.23         | 1.41        |

Table 1 Representative values of available water content (AWC), saturated hydraulic conductivity (Ks), and bulk density (BD) of the 28 soil groups that occur in sugarcane farmlands in Khon Kaen, Thailand.
2.4 ABC analysis

ABC analysis is often used in inventory control to classify items of inventory into a three-level ranking of importance to decide the level of control and record-keeping to which they should be subjected (Dhoka and Choudary 2013). We used it here to classify the 179 land units into a three-tier system of benefit from providing irrigation. We sorted the results of the benefit calculation by the benefit per unit area (THB ha⁻¹) and accumulated benefit (THB). Then divided the areas into A, B, and C categories. In the general ABC analysis, the ranges are 0-70% in rank A, 70-90% in rank B, and >90% in rank C. However, because calculation results were same in the same land unit in this study, the separation of rank have to be same as the separation of land units. Therefore, as the result, we decided the range of each rank as follows: A, 0.0–69.8%; B, 69.8–89.9%, and C, >89.9%.

2.5 Decision tree

For irrigation to be developed efficiently, priority areas for irrigation development should be determined explicitly and objectively. Areas categorized as A rank in the ABC analysis can be regarded as more efficient for irrigation development than areas ranked B and C. Therefore, we tried to analyse the attributes that distinguished areas in Rank A from those in Rank B or C using the decision tree method.

We applied the C4.5 algorithm (Wei Dai and Wei Ji 2014) to construct the decision tree. In the C4.5 algorithm, the amount of entropy is calculated from

\[ E(S) = -\sum_{i=1}^{n} p_i \log_2 p_i \]  

where \( E(S) \) is the entropy of the data set \( S \), \( n \) is the number of classes, and \( p_i \) is the proportion of the \( i \)th class in \( S \). Then the weighted average entropy \( E(A,S) \) is calculated by

\[ E(A,S) = -\sum_{j=1}^{m} \frac{|S_j|}{|S|} E(S_j) \]  

where \( S_j \) is the subset \( j \), and \( m \) is the number of subsets. The gain \( Gain(A,S) \) is calculated from

\[ Gain(A,S) = E(S) - E(A,S) \]  

The intrinsic information of the split \( IntI(A,S) \) is calculated as follows:

\[ IntI(A,S) = -\sum_{j=1}^{m} \frac{|S_j|}{|S|} \log_2 \frac{|S_j|}{|S|} \]  

The gain ratio \( GR(A,S) \) is calculated as follows:

\[ GR(A,S) = \frac{Gain(A,S)}{IntI(A,S)} \]

We defined areas ranked A as Class 1 (C-1), efficient for irrigation development, and those ranked B or C as Class 2 (C-2), inefficient for irrigation development. We used three attributes as explanatory variables: the weather station, the soil group, and the well capacity. We calculated the GR of each attribute and selected the internal nodes attribute with the largest GR. We stopped splitting when all data in a subset belonged to the same benefit class. In this study, we did not use the number of land units but the area of land units to calculate the entropy and the intrinsic information, because the area of each land unit was widely different.

Figure 5: Simulated sugarcane yield of 82 defined simulation conditions every year from 1980 to 2009 under rainfed (a) and automatic irrigation (b) growing condition in Khon Kaen province, Thailand
3 Results and discussion

3.1 Simulation of the sugarcane yield

The calculated sugarcane yield of the RFC ranged from 92 to 194 t ha\(^{-1}\) yr\(^{-1}\), and the average was 153 t ha\(^{-1}\) yr\(^{-1}\). That of the AUIC ranged from 105 to 199 t ha\(^{-1}\) yr\(^{-1}\), and the
average was 177 t ha⁻¹ yr⁻¹. The average calculated sugarcane yield of the 82 simulation conditions under each condition in each year of the simulation is shown in Figure 5, and that over the 30 years for each simulation condition is shown in Figure 6. The average standard deviation of the calculated annual sugarcane yield over the 30 years was 15.2 t ha⁻¹ for RFC and 12.4 t ha⁻¹ for AUIC (Figure 5). For all simulation conditions, the average standard deviation of the annual yield over the 30-year period was 11.0 t ha⁻¹ for RFC and 3.0 t ha⁻¹ for AUIC (Figure 6). The difference in standard deviation between RFC and AUIC was larger for calculated yield averaged across simulation conditions (Figure 6) than across years (Figure 5). This suggests that irrigation reduced the variation in calculated sugarcane yield caused by differences in rainfall from year to year. Conversely, irrigation was less effective at reducing differences in yield caused by soil type.

The average irrigation water used across the 82 simulation conditions in each year is shown in Figure 7 and that for the 30 years in each simulation condition is shown in Figure 8. The average standard deviation in Figure 7 was 63 mm yr⁻¹ and that in Figure 8 was 142 mm yr⁻¹. Thus, the temporal variability was considerably larger than the spatial variability.

### 3.2 ABC analysis

Figure 9 shows the spatial distribution of land units by ABC ranking as well as the land units with negative benefit. All land units with a negative benefit were in the area of well capacity WC-1. Moreover, many of these areas belonged to weather stations WS-2 and WS-4. For these weather stations, the calculated sugarcane yield was large for RFC, so there was little difference between the yields for RFC and AUIC, and the benefit was less than the cost of irrigation.

Figure 10 shows the accumulated benefit of land units plotted against their accumulated percentage of the total area, with land units sorted in order of largest to smallest benefit rank (THB ha⁻¹). The plot is divided to show the contributions of the A, B, and C benefit ranks. We found that 69.8% of the total benefit (rank A land units) is achieved by irrigation development of just 46.7% of the area. The next 20.1% of benefit requires 22.6% of the area, and the final 10.1% of benefit requires 30.7% of the area. The efficiency of irrigation development, calculated by the proportion of the benefit divided by the proportion of the

![Figure 9: Classification of the sugarcane growing fields in Khon Kaen province, Thailand, into ranks by ABC analysis of the estimated benefit of irrigation development. Benefit was estimated from yield modelling, sugarcane prices, and irrigation costs](image)

![Figure 10: Accumulated area–benefit curve from the estimated benefit derived from irrigation development of sugarcane fields in Khon Kaen province, Thailand](image)

| WC-1 | WS  | 0.042 | 0.013 | 1.351 | 0.009 |
| SG  | 0.025 | 0.030 | 2.505 | 0.012 |
| WC-6 | WS  | 0.385 | 0.276 | 1.538 | 0.180 |
| SG  | 0.397 | 0.264 | 2.984 | 0.088 |
| WC-15 | WS  | 0.060 | 0.015 | 1.232 | 0.012 |
| SG  | 0.000 | 0.075 | 2.750 | 0.027 |

Table 3 Results of a decision tree analysis of estimated irrigation benefit for 179 defined land units in Khon Kaen, Thailand. Classification variables are weather station (WS), soil group (SG), and well capacity (WC)

| WC-1 | WS  | 0.042 | 0.013 | 1.351 | 0.009 |
| SG  | 0.025 | 0.030 | 2.505 | 0.012 |
| WC-6 | WS  | 0.385 | 0.276 | 1.538 | 0.180 |
| SG  | 0.397 | 0.264 | 2.984 | 0.088 |
| WC-15 | WS  | 0.060 | 0.015 | 1.232 | 0.012 |
| SG  | 0.000 | 0.075 | 2.750 | 0.027 |

Table 4 Results of a decision tree analysis of estimated irrigation benefit for land units in each well capacity class (WC) (see Table 4 for further information)
area, was 1.49 in the rank A land units, 0.89 in the rank B land units, and 0.33 in the rank C land units.

### 3.3 Decision tree

Recall that from the ABC analysis, we classified rank A land units as class C-1 (efficient for irrigation development) and rank B and C land units as class C-2 (inefficient for irrigation development). The complete classification tree is shown in Figure 11.

For the first internal nodes attribute after the root, well capacity (WC) gave the lowest entropy and had the largest gain and gain ratio (Table 3). Therefore, we chose well capacity (WC) as the first internal nodes attribute. All land units in WC-25 belonged to C-1. From this result, we confirmed that the variable cost in of irrigation (Eq. 5) strongly influenced the benefit.

To identify the next internal nodes attribute, we calculated the entropy, gain, and gain ratio for all land units in WC-1, -6, and -15 by soil group and weather station (Table 4). In WC-1 and WC-15, the soil group had the largest gain ratio so was chosen as the second internal nodes attribute. In WC-6, the weather station had the largest gain ratio so was chosen as the second internal nodes attribute. In WC-1 and WC-6, there were terminal nodes of C-1 and C-2, and the internal node. In WC-15, all land units were in terminal nodes. Then, we tried to classify land units in internal nodes by weather station in WC-1 and by soil group in WC-6. As the result, all land units were in terminal nodes.

We examined characteristics of classification by examining some terminal nodes (Figure 11). Examining the distribution of land units in C-1 and C-2 in WC-6 by weather station, all land units in WS-3 and WS-7 were in C-1, and all those in WS-2 were in C-2, so these formed the next terminal nodes. Weather stations WS-3 and WS-7 had low and WS-2 high average rainfall (Table 2). Thus, the benefit from irrigation development was large in areas with low rainfall and small in areas with high rainfall. Examining terminal nodes in the branch of well capacity WC-6, most soil groups were in C-1. The soil groups in C-2 were SG-17, -36, and -41 needed further analysis, and all other soil groups formed a C-1 terminal node.

For the remaining three soil groups, SG-17, -36, and -41, the C-1 and C-2 terminal nodes were defined by weather station, with WS-1 and WS-6 forming the C-2 terminal node and WS-4 and WS-5 forming the C-1 terminal node.

We examined the yield, rainfall, irrigation, and benefit data for these three soil groups (Table 5). Among these data, we see that the yields under irrigated condition are almost the same within the same soil group. However, under rainfed condition, yields were more varied within soil group, and the land units with larger rainfed yields had smaller yield gaps and lower benefit, and were therefore classified as C-2. The land units with lower rainfed yields showed a greater yield gap and larger benefit and were therefore clas-
At the lowest level of the classification, the classification as C-1 or C-2 appeared to depend on the rainfall of each weather station. However, even though the average rainfall of WS-1, -4, and -6 were almost the same, WS-1 and WS-6 were classified as C-2 and WS-4 was classified as C-1. The standard deviations of annual rainfall at WS-1, -4, and -6 were 177, 234, and 177, respectively (Table 2). The annual rainfall, and therefore yield under rainfed condition, varied widely for WS-4 resulting in the average yield being less than those of the other two weather stations.

Then, we extracted the results of the simulation for the C-2 soil groups (SG-24, -38, -44) occurring in the areas of WS-4 and WS-5 (Table 6). The available water content of SG-24 and SG-44 was low, whereas that of SG-38 was large. When the available water content was low, the yield gap was small because the irrigated yield was low. On the other hand, when the available water content was large, the gap was small because the rainfed yield was large. From these data we conclude that the effect of irrigation development is low when water holding capacity is either low or high.

Examining the distribution by area of C-1 and C-2 between the four well capacities in Figure 11, most of the area in WC-15 and WC-25 belonged to C-1, and most of the area in WC-1 belonged to C-2. From this result, we concluded that irrigation development could be conducted in areas of WC-15 and WC-25, but not in areas of WC-1.

### 4 Conclusions

From this study, we conclude the following.

From the ABC analysis, 70% of the potential benefit could be carried out by the irrigation development of 47% of the cultivation area.

In the decision tree analysis, we confirmed that well capacity was the attribute that most influenced the efficiency of irrigation development. Irrigation development was efficient in the areas with high well capacity (>15 m³ hr⁻¹), but was clearly inefficient where well capacity was low (1 m³ hr⁻¹). At intermediate well capacity, irrigation development was efficient in the majority of the area (18,704 ha efficient versus and 3868 ha inefficient). In areas with higher rainfall, the yield under rainfed conditions was large so the benefit of irrigation development tended to be low. Moreover, soil groups with either low or high available water content resulted in low efficiency of irrigation development.

### References

[1] DGR (2014): Thailand’s Groundwater Database, Depart-
ment of Groundwater Resources. [CDROM].

[2] Dhoka, D., and Choudary, Y.L. (2013): ABC Classification for Inventory Optimization, IOSR Journal of Business and Management, 15, pp.38-41.

[3] Hoogenboom G., J.W. Jones, Wikens, P.W., Porter, C.H., Boote, K.J., Hunt, L.A., Singh, U., Lizaso, J.L., White, J.W., Uryasev, O., Royce, F.S., Ogoshi, R., Gijsman, A.J., and Tsuji, G.Y. (2011). Decision Support System for Agrotechnology Transfer, Version 4.5 [CD-ROM] University of Hawaii, Honolulu, HI, USA.

[4] Inman-Bamber, N.G., Muchow, R.C., and Robertson, M.J. (2002): Dry matter partitioning of sugarcane in Australia and South Africa, Field Crops Res, 76, pp.71-84.

[5] Jintrawet, A., Jongkaewwattana, S., and Inman-Bamber, N.G. (1997): ThaiCane 1.0 simulation model, Final Report of the Development and Validation of Sugarcane Model in Thailand, Multiple Cropping Center, Chiang Mai University (Chiang Mai), pp.102-112. (in Thai).

[6] Kapetch, P., Sakai, K., Pisanjaroen, K., Sansayawichai, T., Cho, T., Nakamura, S., and Nakandakari, T. (2016): Calibration and Validation of Two Crop Models for Estimating Sugarcane Yield in Northeast Thailand, Tropical Agriculture and Development, 60(1), pp.31-39.

[7] LDD (2000): Soil database system, Soil View version 2.0, Land Development Department, Ministry of Agriculture and Cooperatives, Bangkok, Thailand.

[8] OCSB (2013): Report of Sugarcane Growing Area 2012/13, Office of Cane and Sugar Board, Bangkok, 125p. (in Thai).

[9] OCSB (2015): Report of Sugarcane Growing Area 2014/15, Office of Cane and Sugar Board, Bangkok, 123p. (in Thai).

[10] OCSB, and KKU (2015): Report of the cost for sugarcane production in 2014/2015, Khon Kaen Printed, Khon Kaen, 86p. (in Thai).

[11] O’Leary G.J. (2000): A review of three sugarcane simulation models with respect to their prediction of sucrose yield, Field Crops Res, 68, pp.97-111.

[12] Singels A., and Bezuidenhout, C.N. (2002): A new method of simulating dry matter partitioning in the CANEGRO sugarcane model, Field Crops Res, 78, pp.151-164.

[13] Singels A., Donaldson, R.A., and Smit, M.A. (2005): Improving biomass production and partitioning to sugarcane: theory and practice, Field Crops Res, 92, pp.291-303.

[14] TMD (2010): Average Data of Rainfall in Thailand from 1979-2008, Meteorological Development Bureau, Thai Meteorological Department. [CDROM].

[15] Dai, W., and Ji, W. (2014): A MapReduce Implementation of C4.5 Decision Tree Algorithm, International Journal of Database Theory and Application, 7 (1), pp. 49-60.

Discussion open until June 30, 2017