The optimal harvest decisions for natural and artificial maturation mangoes under

uncertain demand, yields, and prices

Sheng-I Chen[a,b], Wei-Fu Chen[a,b]

[a]Department of Industrial Engineering and Management, College of Management,

National Yang Ming Chiao Tung University

[b]Department of Industrial Engineering and Management, College of Management,

National Chiao Tung University

1001 University Rd.

Hsinchu, 30010, Taiwan

Corresponding Author:

Sheng-I Chen, Associate Professor

Phone: +886-3-5131454

Email: sichen@nctu.edu.tw

**Abstract**

This study focuses on the decisions of picking, inventory, ripening, delivering, and selling mangoes in a harvesting season. Demand, supply, and prices are uncertain, and their probability density functions are fitted based on actual trading data collected from the largest spot market in Taiwan. A stochastic programming model is formulated to minimize the expected cost under the considerations of labor, storage space, shelf life, and transportation restrictions. We implement the sample-average approximation to obtain a high-quality solution of the stochastic program. The analysis compares deterministic and stochastic solutions to assess the uncertain effect on the harvest decisions. Finally, the optimal harvest schedule of each mango type is suggested based on the stochastic program solution.

**Keywords:** Fresh agricultural products, harvest schedule, stochastic programming, sample-average approximation

1.    **Introduction**

This study considers a sequence of decisions for harvesting mangoes, where the fruit farmer determines the volume and timing for picking, storing, transporting, and selling different types of mangoes. These decisions are usually made on a daily basis but need to adjust frequently because of the fluctuated price and imbalances between demand and supply. Additionally, mango shelf-live are relatively shorter than other fresh fruits. Once picking the mangoes, they must be sold out within a day or two. Finally, the harvest season is very short, and there is often a shortage of personnel throughout the season. We tackle these challenges with the considerations of labor availabilities, transporter capacity, and storage spaces. The objective is to maximize the overall profit for multiple mango varieties during a harvest season. A *stochastic program* (SP) is formulated with the considerations of uncertain yield, demand, and prices. The distributions of uncertain parameters are estimated by using historical data collected from a spot market.

The processes for harvesting mangoes are depicted in Fig. 1. The first type is the naturally-matured mangoes, which include Mangifera indica Linn and Irwin. The harvesting operations start from picking, stocking, delivering to trading. The shelf life of naturally-matured mangoes is two days. After picking, mangoes can be shipped to the market directly or stocking at the warehouse and then selling on another day. The second type is artificially-matured mangoes, including Jin-Hwang, Yu-Wen, Sensation, and Keitt. These mangoes must be harvested before they fully matured. The calcium carbide is applied to reduce the maturation time, and the average duration of the artificial maturation process is three days [1]. Once the artificial maturation operation is complete, mangoes need to be kept for another day to ensure that their skin color and taste are ready for consumers.
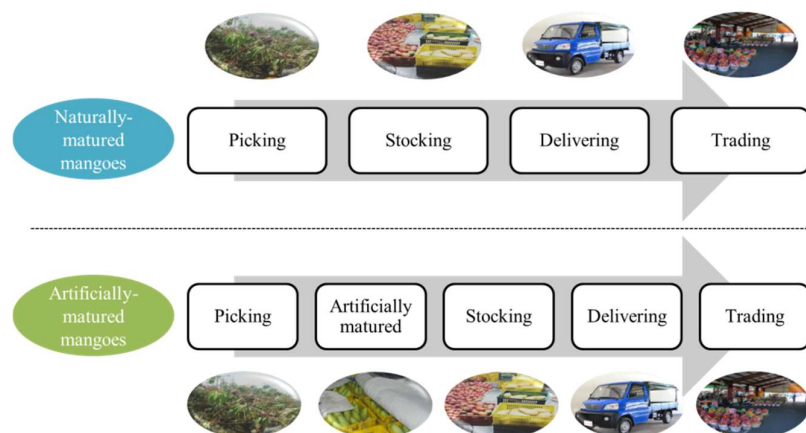


**Fig. 1.** The harvesting processes of naturally- and artificially-matured mangoes.

There are different ways to trade mangoes. The first mode is to sell mangoes in a spot market, where buyers negotiate mangoes prices with farmers according to mango qualities and overall supply quantities in the market. The transaction is confirmed whenever the farmer accepts the offer. Since mango demand and supply are highly fluctuated, the selling prices usually vary from day to day. The second mode to trade mangoes is the price contract, in which the wholesaler and farmer negotiate a fixed price prior to the harvest season. The contracted price can reduce price variation but less favorable for farmers in Taiwan. This is because the mango qualities differ between different ranches and the contracted price is usually much lower than the selling price in the spot market.

Prior work has considered uncertain factors for making agricultural decisions. For example, scholars analyzed land allocation decisions under uncertain demand [2]. They developed a stochastic programming model with the objective to maximize the probability of satisfying demand. There is a strong correlation between demand and prices for selling mangoes in a spot market. Mango price tends to be lower on the day of higher quantity and higher when the trading volume is low. **Fig. 2** illustrates the selling price and trading quantity on each day for different mango varieties during the harvest season. We found that few papers have developed models to cooperate the inter-effect between demand and supply. This study uses the actual data collected from [3] the Agriculture and Food Agency [3] along with forecasting models to construct the relationship between demand and prices.
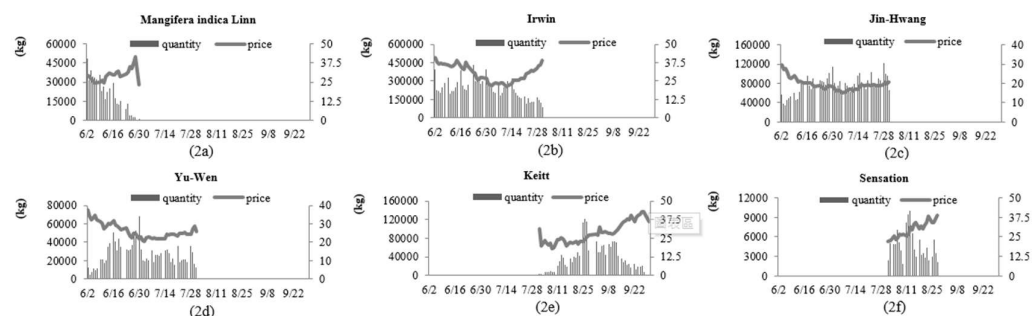


**Fig. 2** Trading volumes and selling prices for each mango variety in the spot market.

During the harvesting season, mango farmers determine how many mangoes to be picked on each day, when to start the artificial maturation operation, and how many mangoes to be sold. These decisions are usually made according to farmers' experiences. To tackle the increasing uncertainties of price and yield, scholars suggested risk-sharing strategies, such as insurances [4]. Since the conclusion is based on a questionnaire survey result among farmers in Dutch, these suggested strategies may not be available in our case. In the area of mathematical models, a survey paper

highlighted the applications of stochastic programming in supply chain planning under uncertain demands [5]. Other fields applications, such as finance, manufacturing, telecommunication, transportation, and energy can be found in another review paper [6]. Among these applications, a common assumption of stochastic programming model is that the uncertain parameter's distribution function must be given in order to formulate the *deterministic equivalent program* and thus the problem becomes computationally trackable. Interested readers may refer to the textbook for a comprehensive review of stochastic programming models and solution approaches for solving each of them [7]. Another approach to model uncertain considerations is robust optimization. Such a model assumes that the variances of uncertain parameters are bounded. Instead of considering the expected performance, the goal is to find the optimal solution considering the worst-case scenario [8]. The drawback of robust optimization is that its decisions may be over-conservative when parameter values are high variance. An application of robust optimization models in agriculture may refer to the wine grape harvest scheduling problem done by Bohle *et al.* [9].

This study applies the stochastic programming model to explore the operational decisions for mango farmers. Both demand and production yields are predicted via the *bass models*. The forecast error of each mango type is assumed normally distributed, where mean and variance are estimated according to the forecast errors. Additionally, we assume that the forecast error distribution is identical and independent on each day during the harvesting season. Uncertain parameters are constructed based on the time-series forecast and forecast error distribution. The stochastic programming model's objective function minimize the expected cost and constraints capture the precedent operations in a harvesting process as well as resource capacities.

## 2.    Literature review

Harvesting and processing planning have been studied broadly during past decades. An early work analyzed the production schedule for a fresh tomato packing house [10]. Higgins developed models to determine a harvesting schedule to minimize the variability of the daily supply of sugar cane to the mill and variability in daily transportation resource usage [11]. Caixeta-Filho *et al.* considered the production plan of lily flower that maximizes net revenue for the farm by taking into account the planting week and expected harvest week [12]. Another work by Caixeta-Filho focused on the orange harvesting scheduling in the juice processing industry [13].

Recently, Vizvári *et al.* analyzed crop planting decisions with the development of

stochastic programs [14]. The crop yield was assumed to be uncertain and cyclic due to the crop's memory of drought. A chance constraint was formulated to ensure that a certain level of demand is satisfied. Their results suggested that planting crops in a smaller area can provide an extra supply to cover requirements during the low yield season. Another related work studied the crop rotation decision where both yield and demand are fluctuated and uncertain [2]. The objective function is to maximize the probability of demand satisfaction. In the livestock industry, Guan and Philpot analyzed the milk distribution decision in an arborescent supply chain. They developed models to predict milk supplies and used random forecast errors for modeling uncertain parameters in the multi-stage stochastic program [15]. Other researchers assessed the productivity risk under different levels of price settings. A dynamic stochastic programming model was proposed to evaluate both the long-term and short-term rotation strategies for planting crops [16].

Recently, attention has turned to the question about freshness and shelf life effects on agricultural decisions. Ahumada and Villalobos analyzed both harvesting and inventory decisions simultaneously with the consideration of shelf life, where the selling price was assumed to decrease over time for the decaying in the quality of agriculture products [17,18]. The objective maximized the profit under labor force capacity limits. A related work by Ferrer *et al.* developed a MIP model for grape planting decisions. Their model determined laborer and machine allocations and the sequence of fields to be harvested [19]. The grape quality was formulated as a non-increasing function in time. Similar considerations were studied in another paper by Zhang and Wilhelm [20]. In the area of inventory models, Lodree and Uzochukwu developed a two-period inventory model for fresh products to consider product deterioration, uncertain demand, lead times, and consumer preferences [21]. Also, Noparumpa *et al.* developed models to determine production allocation decisions for a winemaker. They considered both the expected profit and risks of quality rating degradation [22].

We found that prior studies focused on specific operational decisions in the agriculture industry. Yet, none of them have considered the end-to-end process. This study aims to coalesce data analysis and decision model to assess interactions between decisions in different operations. Furthermore, the agriculture product price is strongly correlated with yields. The difference between this study and related literature is that the relationship between prices and yields is considered in the harvesting decision. Finally, uncertainties significantly affect agriculture decisions. This study develops stochastic programming model of mango harvesting that

explicitly considered uncertain demand, yields, and prices. The next section will explain the detailed procedure of modeling uncertain parameters and the mathematical formulas in stochastic programming.

## 3.    Methods

The harvesting planning comprises a sequence of operational decisions for different varieties of mangoes from June to September. To model uncertain yields and demand, we first apply forecasting the *Bass diffusion model* and use time-series trading data in previous years to estimate model parameters. Forecasting errors are calculated using another data set of the testing data. For each mango type, the forecasting error is assumed to be an independent and identical normal distribution, where its mean value and variance are estimated using forecasting errors obtained beforehand. To model the selling price, a regression model is used to explore the relationship between the price and demand. The contents of this section firstly describe key features and parameters for each mango variety. Then, we explain the data collection approach and how the uncertain parameters are setup using fitting results. The last subsection presents the stochastic programming model.

### 3.1 Data and assumptions

This study considers the common types of mangoes in Taiwan, and each of them has a different maturation process, picking schedule, and shelf life. The naturally-matured mangoes (including Mangifera indica Linn and Irwin) have a short shelf life of about two days, while the artificially-matured mangoes (including Jin-Hwang, Yu-Wen, Sensation, and Keitt) have a longer shelf life of up to thirty days. The harvest timing varies from year to year. **Fig. 3** displays the sales data collected from *Tainan County* between June and September in 2015. The line chart shows the weekly sales volume for each mango type, and the pie chart depicts the proportion of sales volume. The harvest season starts in June and ends by the end of September. The highest volume occurs on the fourth week of June. Also, most trading takes place in June and July, and August and September have less than one-fifth of the entire volume. *Irwin* has the most trading volume among the six mango types, which accounts for more than 60% of the overall volume. In comparison, *Jin-Hwang* is the second-highest product with a 20% market share.
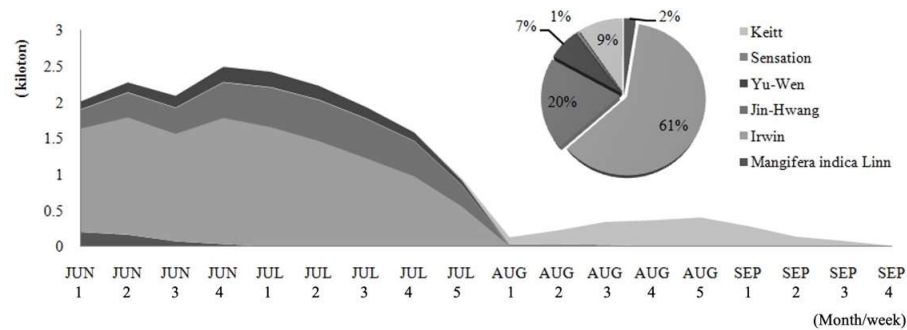
**Fig. 3.** The weekly trading volume and market share of each mango type during a harvesting season.

**Table 1** summarizes the major characteristics of mangoes considered in this study. Mangifera indica Linn and Irwin are natural-maturation mangoes. The other mango types must perform the artificial maturation, and this process will take three days. The seasons for picking Mangifera indica Linn start from the beginning of June until the end of July, and for Irwin, Jin-Hwang, and Yu-Wen are between June and August. The Sensation and Keitt picking seasons are between August and September.

**Table 1.** Maturing types and picking time windows for different mango varieties.

| Mango varieties | Is artificial maturation required? | Picking timeline |
|---|---|---|
| Mangifera indica Linn | No | June-July |
| Irwin | No | June-August |
| Jin-Hwang | Yes | June-August |
| Yu-Wen | Yes | June-August |
| Sensation | Yes | August-October |
| Keitt | Yes | August-October |

### 3.2 Time-series forecasts and forecast error analysis

This section describes how the distributions of uncertain parameters for the stochastic programming model are constructed. Mango demand will grow rapidly with the increase in productivity at the beginning of the season. Once the peak is passed, the demand will begin to shrink exponentially. Such a pattern can well fit by the use of the *Bass diffusion model* [23]. We segregate trading data by mango types and use each of them to estimate model parameters separately.

Let $f(t)$ be the density of adopters at time period $t$ with an associated cumulative distribution $F(t)$. Notation $p$ represents the coefficient of innovation (or external influence) and $q$ is the coefficient of imitation (or internal influence). Thus, the

proportion of demand not being at time period $t$ is equal to the portion of new innovator and new imitator as the following equation $f(t)/(1 - F(t)) = p + qF(t)$. We further denote $m$ as the total demand of adopters, $n(t)$ as the number of adopters at time $t$ defined as $n(t) = mf(t)$, and the cumulative demand of adopters at period $t$ as $N(t)$. Hence, the number of new adopters can be determined by $n(t) = (p + q(N(t)/m))(m - N(t))$. Let $Y_t$ be the cumulative adopters and $S_t$ be the new ones observed at time $t$. The estimation of new adopters is as the following equation: $S_t = m\hat{p} + (\hat{q} - \hat{p})Y_{t-1} - (\hat{q}/m)Y^2_{t-1} + \hat{\varepsilon}(t)$. The objective of the estimation is to find the parameter values of $p$ and $q$ such that the total error is minimal. The actual trading volume is aggregated in a weekly basis for fitting the forecasting models.

**Table 2** shows the estimated parameters in the Bass model for each mango variety. As a result, all $\hat{q}/\hat{p}$ ratios are greater than one (the smallest value is 2.0 for Mangifera indica Linn and the largest ratio is 28.1 for Keitt). It implies that the demand increas or decreases exponentially over time. Also, the sales volume in the previous time period has positive effects (i.e., $(\hat{q} - \hat{p})Y_{t-1} > 0$).

**Table 1.** The fitting results of internal influence and external influence factors

| Mango variety | External influence $(\hat{p})$ | Internal influence $(\hat{q})$ | $\hat{q}/\hat{p}$ |
|---|---|---|---|
| Mangifera indica Linn | 0.0500 | 0.1000 | 2.0 |
| Irwin | 0.0134 | 0.0394 | 2.9 |
| Jin-Hwang | 0.0064 | 0.0495 | 7.7 |
| Yu- Wen | 0.0106 | 0.0578 | 5.5 |
| Sensation | 0.0324 | 0.1230 | 3.8 |
| Keitt | 0.0042 | 0.1180 | 28.1 |

The Bass models perform well in capturing demand trends during the harvesting season. **Fig. 4** shows the actual and forecast trading volumes for different mango varieties. As one can tell, the model can precisely predict the start time, peak time, and end time of each mango type. Except for Jin-Hwang**,** Sensation, and Keitt mangoes, the predicted peak volume is very close to the actual situation. We further analyze the accuracy of the Bass model. **Table 3** shows the *Mean Absolute Percentage Error* (MAPE) of each mango type, where MAPE is defined as the ratio of the expectation of absolute errors divided by actual demand $\frac{1}{n}\sum_{t=1}^{n}\frac{|actual_t - forecas_t|}{actual_t} \times 100\%$. The MAPEs of Mangifera indica Linn, Jin-Hwang, and Yu-Wen are less than 20%, Irwin and Keitt are about 30%, and Sensation
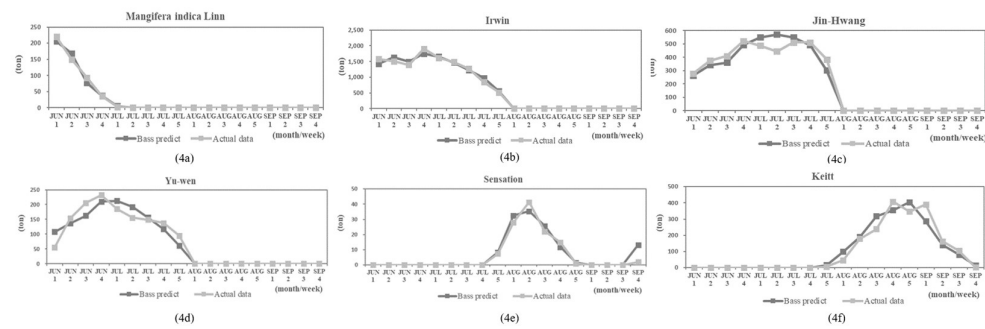
performed the worst at 40%.



**Fig. 4.** Comparing the trading volumes of actual and forecast in a time series.

**Table 2.** The Bass model MAPE performances.

| | Mango variety | | | | | |
|---|---|---|---|---|---|---|
| | Mangifera indica Linn | Irwin | Jin-Hwang | Yu- Wen | Sensation | Keitt |
| MAPE | 18% | 27% | 14% | 13% | 40% | 30% |

### 3.3 Stochastic programming model

Stochastic programming has been widely applied for analyzing the optimal decision under uncertain environments [6]. In this study, demand, yield, and price are considered as random variables, the probability distribution of which has been explained in the foregoing subsection. The objective is to minimize the expected cost, which includes inventory cost, shipping cost, artificially maturing cost, and sales revenue across all mango varieties during the multi-month harvest season. The decision variables involve picking quantity, inventorying quantity, artificial maturation quantity, shipping quantity, and sales volume on each time period. **Table 4** summarizes notations used in the stochastic programming model.

**Table 4.** Notations.

***Sets:***

| | |
|---|---|
| $I$ | The set of naturally-matured mangoes |
| $J$ | The set of artificially-matured mangoes |
| $K$ | The set of all mango varieties, $K = I \cup J$ |
| $\Omega$ | The set of all random scenarios, $\omega \in \Omega$ |
| $T$ | The set of planning time periods |

***Decision variables:***

| | |
|---|---|
| $x_{kt}(\omega)$ | The number of $k$ type mangoes picked at time $t$ in scenario $\omega$ |
| $y_{it}(\omega)$ | The sales volume of naturally-matured mangoes at time $t$ in scenario $\omega$ |
| $y_{jt}^1(\omega)$ | The sales volume of one-day-old artificially-matured mangoes |
| $y_{jt}^2(\omega)$ | The sales volume of two-day-old artificially-matured mangoes |
| $r_{it}(\omega)$ | The shipping quantity of naturally-matured mangoes |
| $r_{jt}^1(\omega)$ | The shipping quantity of one-day-old artificially-matured mangoes |
| $r_{jt}^2(\omega)$ | The shipping quantity of two-day-old artificially-matured mangoes |
| $i_{it}(\omega)$ | The number of naturally-matured mangoes |
| $i_{jt}^w(\omega)$ | The inventory level of mangoes waiting for artificial maturation |
| $i_{jt}^m(\omega)$ | The inventory level of artificially-matured mangoes |
| $w_{jt}(\omega)$ | The number of mango $j$ ripened artificially at time $t$ in scenario $\omega$ |
| $h_{jt}(\omega)$ | The number of unsold one-day-old artificially-matured mangoes |
| $a_{kt}(\omega)$ | The number of mangoes not yet picked |

***Parameters:***

| | |
|---|---|
| $p_{kt}(\omega)$ | The selling price of mango type $k$ at time period $t$ in scenario $\omega$ |
| $v_{it}$ | The inventory cost of naturally-matured mangoes |
| $v_{jt}^w$ | The inventory cost of mangoes waiting for artificial maturation |
| $v_{jt}^m$ | The inventory cost of artificially-matured mangoes |
| $\beta_{kt}(\omega)$ | The yield of mango type $k$ at time period $t$ in scenario $\omega$ |
| $d_{kt}(\omega)$ | The demand of mango type $k$ at time period $t$ in scenario $\omega$ |
| $O^N$ | Maximum inventory level for naturally-matured mangoes |
| $O^A$ | Maximum inventory level for artificially-matured mangoes |
| $C$ | The capacity limit of picking mangoes |
| $l$ | The transportation cost |

**Fig. 5** illustrates our ideas to formulate the MIP model, where the flow chart at the top represents the harvesting process for naturally-matured mangoes, and another one is for artificially-matured mangoes. Each rectangular block represents either picking, artificially maturing, shipping, or trading operation in both flow charts, while trapezoid blocks represent stocking mangoes with various statuses. The directed edge represents the sequence to perform two consecutive operations. Most blocks are connected by less than or equal to one incoming and out-going arcs, except for the operations of picking and stocking matured-and-unsold artificially-matured mangoes. Since the farmer can defer artificially maturing decisions, two out-going arcs depart from the picking operation. Another block with multiple out-going arcs is the matured-and-unsold stocking, in which the solid out-going arc represents the first attempt of farmers to deliver and sell the one-day-old mangoes, and the dashed out-going arc is the second attempt for dealing with the two-day-old mangoes. Additionally, two in-coming arcs connect to the matured and unsold stocks, where the sold arc represents the influx of one-day-old mango, and the dashed arc is the returning unsold mangoes.



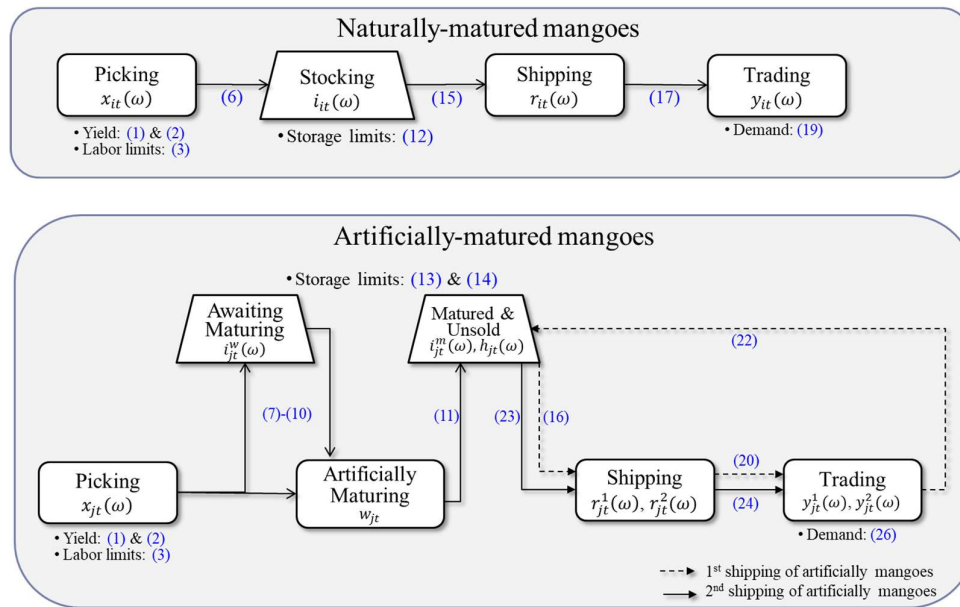**Fig. 5.** Illustrating the relationship between constraints and decision variables in the proposed stochastic programming model.

The stochastic programming model of mango harvest and distribution is as the following:

$$min\, E\left[\sum_{t\in T}\sum_{i\in I}v_{it}\,i_{it}(\omega)+\sum_{t\in T}\sum_{i\in I}l\,r_{it}(\omega)+\sum_{t\in T}\sum_{j\in J}\left(v_{jt}^{w}\left(i_{jt}^{w}(\omega)+\right.\right.\right.$$

$$w_{jt}(\omega)) + v_{jt}^m \, i_{jt}^m(\omega) + l \, r_{jt}^1(\omega) + v_{jt}^m \, h_{jt}(\omega) + 3l \, r_{jt}^2(\omega)) - \sum_{t \in T} \sum_{i \in I} p_{it} \, y_{it}(\omega) -$$

$$\sum_{t \in T} \sum_{j \in J} p_{jt} \left( y_{jt}^1(\omega) + y_{jt}^2(\omega) \right) \Big],$$

*Subject to*

$$x_{kt}(\omega) \le \beta_{kt}(\omega) \qquad\qquad \forall t \in \{1\}, \forall k \in K, \qquad (1)$$

$$x_{kt}(\omega) \le \beta_{kt}(\omega) + a_{kt-1}(\omega) \qquad\qquad \forall t \in \{2..|T|\}, \forall k \in K, \quad (2)$$

$$\sum_{k \in K} x_{kt}(\omega) \le C \qquad\qquad \forall t \in T, \qquad (3)$$

$$\beta_{kt}(\omega) - x_{kt}(\omega) = a_{kt}(\omega) \qquad\qquad \forall t \in T, \forall k \in I \cup J, \qquad (4)$$

$$i_{it}(\omega) = x_{it}(\omega) \qquad\qquad \forall t \in T, \forall i \in I, \qquad (6)$$

$$i_{jt}^w(\omega) = x_{jt}(\omega) - w_{jt} \qquad\qquad \forall t \in \{1\}, \forall j \in J, \qquad (7)$$

$$i_{jt}^w(\omega) = i_{jt-1}^w(\omega) + x_{jt}(\omega) - w_{jt} \qquad\qquad \forall t = \{2 \ldots |T|\}, \forall j \in J, \quad (8)$$

$$w_{jt}(\omega) \le x_{jt}(\omega) \qquad\qquad \forall t \in \{1\}, \forall j \in J, \qquad (9)$$

$$w_{jt}(\omega) \le i_{jt-1}^w(\omega) + x_{jt}(\omega) \qquad\qquad \forall t = \{2 \ldots |T|\}, \forall j \in J, \quad (10)$$

$$i_{jt}^m(\omega) = 0 \qquad\qquad \forall t = \{1 \ldots 3\}, \forall j \in J, \qquad (11)$$

$$i_{jt}^m(\omega) = w_{jt-3}(\omega) + i_{jt-1}^m(\omega) - \; r_{jt}^1(\omega) \qquad\qquad \forall t = \{4 \ldots |T|\}, \forall j \in J, \quad (12)$$

$$\sum_{\forall i \in I} i_{it}(\omega) \le O^N \qquad\qquad \forall t \in T, \qquad (13)$$

$$\sum_{\forall j \in J} (i_{jt}^w(\omega) + \sum_{s=1}^t w_{js}) \le \; O^A \qquad\qquad \forall t = \{1 \ldots 3\}, \qquad (14)$$

$$\sum_{\forall j \in J} ( i_{jt}^w + \sum_{t-2}^t w_{jt}(\omega) + \; i_{jt}^m - h_{jt}(\omega)) \le O^A \qquad\qquad \forall t = \{4 \ldots |T|\}, \qquad (15)$$

$$r_{it}(\omega) \le i_{it-1}(\omega) \qquad\qquad \forall t = \{2 \ldots |T|\}, \forall i \in I, \quad (16)$$

$$r_{jt}^1(\omega) \le i_{jt-1}^m(\omega) \qquad\qquad \forall t = \{5 \ldots |T|\}, \forall j \in J, \quad (17)$$

$$y_{it}(\omega) \le r_{it}(\omega) \qquad\qquad \forall t = \{2\dots|T|\}, \forall i \in I, \quad (18)$$

$$y_{it}(\omega) \le d_{it}(\omega) \qquad\qquad \forall t = \{2\dots|T|\}, \forall i \in I, \quad (19)$$

$$y_{jt}^1(\omega) \le r_{jt}^1(\omega) \qquad\qquad \forall t = \{5\dots|T|\}, \forall j \in J, \quad (20)$$

$$h_{jt}(\omega) = 0 \qquad\qquad \forall t = \{1\dots3\}, \forall j \in J, \quad (21)$$

$$h_{jt}(\omega) = r_{jt}^1(\omega) - y_{jt}^1(\omega) \qquad\qquad \forall t = \{4\dots|T|\}, \forall j \in J, \quad (22)$$

$$r_{jt}^2(\omega) = 0 \qquad\qquad \forall t = \{1\dots4\}, \forall j \in J, \quad (23)$$

$$r_{jt}^2(\omega) \le h_{jt-1}(\omega) \qquad\qquad \forall t = \{5\dots|T|\}, \forall j \in J, \quad (24)$$

$$y_{jt}^2(\omega) = 0 \qquad\qquad \forall t = \{1\dots4\}, \forall j \in J, \quad (25)$$

$$y_{jt}^2(\omega) \le r_{jt}^2(\omega) \qquad\qquad \forall t = \{5\dots|T|\}, \forall j \in J, \quad (26)$$

$$y_{jt}^1(\omega) + y_{jt+1}^2(\omega) \le d_{jt}(\omega) \qquad\qquad \forall t = \{1\dots|T|\}, \forall j \in J, \quad (27)$$

*All decision variables are nonnegative.* $\qquad\qquad\qquad\qquad\qquad$ (28)

The objective function is to minimize the expected costs, where the first and second terms represent the inventory cost and transportation cost of artificially-matured mangoes, respectively. The third term $v_{jt}^w\left(i_{jt}^w(\omega) + w_{jt}(\omega)\right)$ accounts for the inventory cost of one-day-old artificially-matured mangoes. The fourth term $v_{jt}^m\, i_{jt}^m(\omega)$ represents the inventory cost of two-day-old artificially-matured mangoes. Since the freshness will decay over time, we set a higher inventory cost of artificially-matured mangoes than mangoes waiting artificially-matured (i.e., $v_{jt}^m \gg v_{jt}^w$). The fourth term $l\, r_{jt}^1(\omega)$ accounts for the shipping cost of one-day-old artificially-matured mangoes. The fifth term $v_{jt}^m\, h_{jt}(\omega)$ is the inventory cost of unsold mangoes. These mangoes must return to the farmer and then ship to the market on another day. Also, a third trip may need for returning the unsold two-day-old mangoes, and therefore, the total shipping cost of two-day-old artificially-matured mangoes is modeled as three times of per-unit shipping cost $3l\, r_{jt}^2(\omega)$. The terms $\sum_{t\in T}\sum_{i\in I} p_{it}y_{it}(\omega)$ and $\sum_{t\in T}\sum_{j\in J}(p_{jt}(y_{jt}^1(\omega) + y_{jt}^2(\omega)))$ account for the total sales incomes of selling naturally-matured mangoes and artificially-matured mangoes, respectively.

In Constraints (1) and (2), the picking quantity must be less than or equal to the number of mangoes ready to be harvested. Since the picking decision can delay for one day, the right hand side of constraints (2) includes both yield and unpicked mangoes in the previous time period. Constraint (3) considers the capacity limit of picking mangoes at each time period. Constraint (4) is used to compute the number of unpicked mangoes $a_{kt}(\omega)$, which is equal to yield minus picking quantity. Constraint (6) ensures that the inventory quantity of naturally-matured mangoes $i_{it}$ is equal to the picking quantity $x_{it}$.

The artificially-matured mango can be preserved for a long time before maturity. Constraints (7) and (8) determine the inventory quantity of mangoes waiting for artificial maturation $i_{jt}^w(\omega)$, equal to the previous inventory plus picking quantity $x_{jt}$, and then minus artificially-matured quantity $w_{jt}$. Constraints (9) and (10) ensure that the ripening quantity cannot exceed inventory and picking. Constraints (11) and (12) compute inventory of mangoes that have artificially matured. Since the maturing process takes three days, the inventory levels in the first three time periods are zeros ( $i_{jt}^m(\omega) = 0, \ t = 1..3$ ) . In constraint (12), the $i_{jt}^m(\omega)$ is equal to the artificially-matured quantity three days ago, plus previous inventory, and then minus the shipping quantity of one-day-old matured mangoes $r_{jt}^1(\omega)$.

The artificially and naturally-matured mangoes are kept separately. Constraint (13) ensures that the total inventory of naturally-matured mangoes is less than or equal to the warehouse capacity $O^N$. In constraint (14), the inventory level of artificially matured mangoes in the first three days, equal to mangoes waiting to be matured plus the total amount of ripening from the first day to the present, cannot exceed the maximal capacity $O^A$. Similarly, the left-hand side in constraint (15) determines the inventory levels on the fourth day and after cannot exceed the capacity limit, where the mango inventory is equal to the mangoes waiting to be mature, the mangoes in ripening, the mangoes that have matured, and minus unsold one-day-old mangoes. **Fig. 6** is a numerical example to explain the left-hand-side formulations in constraints (14) and (15). Suppose that seven mangoes are picked on the first day, and two of them are ripening. Therefore, five unripened mangoes and the total inventory at the end of the day are seven. Five mangoes are harvested on the second day, and thus the total inventory is twelve, including the first day's harvest. Additionally, assuming that two mangoes are beginning to ripen, and therefore the total ripening mangoes are four, two of which one day old matured and two are two days old matured. On day five, the inventory level includes mangoes picked from day one to day four, and then minuses two mangoes ripened on day one and sold on day five.
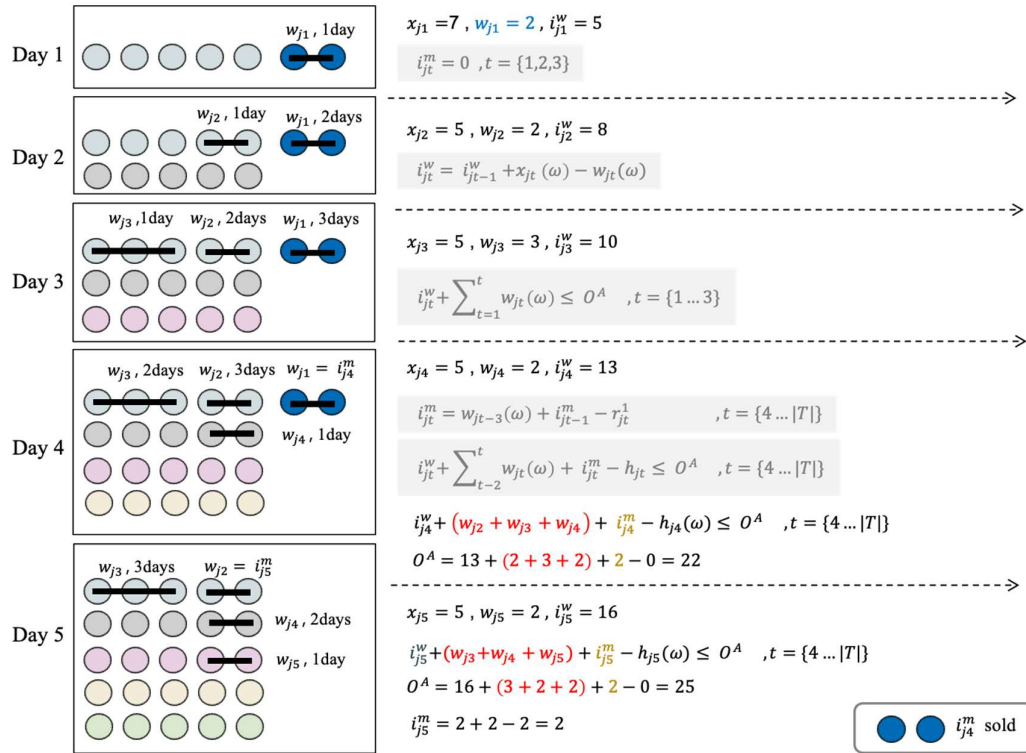
**Fig. 6.** The numerical example to illustrate the inventory levels of artificially-matured mangoes defined at the left-hand-sides in constraints (14) and (15).

Constraint (16) ensures that the shipping quantity of naturally-matured mangoes $r_{it}$ is less than or equal to the inventory in the previous time period to account for lead-time of inventory and transportation operations. Constraint (17) ensures that the quantity to delivery one-day-old artificially-matured mangoes cannot exceed the ripened mangoes. Constraint (18) ensures that the naturally-matured mango selling quantity is less than or equal to the shipping quantity. Constraints (19) ensures the naturally-matured mangoes selling quantities are less than or equal to the demand. Constraints (20) ensure that the selling quantity of one-day-old artificially-matured mangoes are less than or equal to the shipping quantity. Constraints (21) and (22) determine the number of unsold one-day-old mangoes, which is zero in the first three days, and then defined as shipments minus sales. In constraints (23) and (24), the number of two-day-old mangoes is zero in the first four days or no more than stocks after the fourth day. Similarly, constraints (25) and (26) specify the boundary of selling two-day-old mangoes each day. In constraint (27), the total sales of one-day-old and two-day-old mangoes cannot exceed the demand.

# 4.   Results

We use the forecasts and forecast error distributions to construct the probability distribution of the uncertain parameters of the stochastic programming model. The forecast error at each time period follows an independent and identical normal distribution, where the mean and standard deviation are estimated from the testing data set. The demand and yield at each time period are generated using forecasts and forecast error distributions, both of which are sampled from the same probability distribution but sampled separately. In addition, the mango price is generated by a linear function of mango demand, where its parameters are fitted by the same data set used to build the forecast model. It is worth mentioning that the fitting results show an opposite relationship between the sales price and trading volume of all mango types, and this trend is the same as that of most fresh products.

## 4.1    Convergence analysis of SAA

The sampling sizes of uncertain parameters affect both the model accuracy and the runtime to solve stochastic programs. Models using larger sampling size provide a better approximation, but they are usually more difficult to solve. There is no clear cut to determine the best sampling size. We implement the *sample average approximation* (SAA) to obtain a well quality solution in time for the proposed stochastic programming model [6]. The algorithm samples coefficients in both objective function and constraints and then resolves the new sampled problems iteratively. The optimal objective value will stabilize once the sample size increased to a sufficient number. We stop the algorithm when the gap of objective values in two consecutive iterations is less than a tolerated value. **Fig. 7** displays the expected profit (that is, the objective value multiplied by -1) of stochastic programs in different sample sizes. The objective value decreases as the sample size increases, and when the same size increases to forty, it will stabilize at about five million.
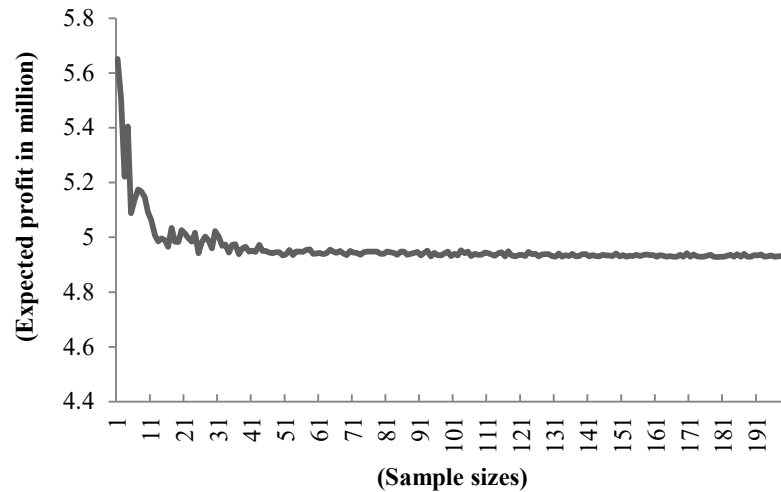
**Fig. 7.** The expected profits of stochastic program solutions with different sample sizes.

In the following analysis, we use a sample size of two hundred. We believe that such a sample size would provide relatively robust results, and similar findings would be concluded even using larger sample sizes.

## 4.2 Comparing deterministic and stochastic solutions

This section compares the deterministic and stochastic models to assess the value of stochastic solutions and understand the impact of uncertain forecast errors on profits. The benchmark model, known as the *expected value* (*EV*) model, is formulated as a deterministic mathematical program using the expected values of uncertain parameters. The *EV* solution is then applied to each sampled problem instance to obtain the *expectation of expected value solution* (*EEV*). For minimization problems, the stochastic program solution is less than or equal to *EEV*. This is because the *EV* solution is suboptimal (or even infeasible) for each sampled problem instance. The *value of stochastic solution* (*VSS*)*,* defined as *EEV – SP,* represents the potential loss of profit for mango farmers using the expected value to make harvest decisions. Another implication is the impact of uncertain forecast errors on profits. Our analysis uses two hundred randomly generated scenarios, and the results are shown in **Table 5**. The objective function is to minimize the total cost. Thus, a smaller objective value means a more profitable solution. The *EV* objective value (approximately minus 5.8 million) is less than *SP*, but when the *EV* solution applies to all other scenarios, the expected objective value is -4,911,662. Therefore, *EV* is about 20% higher than *EEV* (from -5,816,137 to -4,911,662). This implies that the *EV* objective value is too optimistic and may mislead decision-makers. In contrast, if decision-makers adopt the *SP* solution rather than *EV*, they can increase profit by $50,252.

**Table 3.** Comparing expected-value model and stochastic program model solutions

| EV | SP | EEV | VSS |
|---|---|---|---|
| -5,816,137 | -4,961,914 | -4,911,662 | 50,252 |

### 4.3 The optimal harvest planning

The *SP* solution provides abundant information about the timing and quantities of each operational decision. **Table 6** is the color code of each operation that will be used for presenting the optimal harvest planning in **Table 7**. Due to the limitation of manuscript length, **Table 7** only contains the optimal quantity for picking, stocking, delivery, and sales in June. In these tables, each cell shows the lower and upper limits of the best solution in all scenarios. For example, the picking quantities of Mangifera indica Linn on June 5$^{th}$ are between 11 and 13 tons. If there is only a single value in a cell, it means that all scenarios obtain the same solution.

**Table 4.** Color coding for different operations

| Operations | Color |
|---|---|
| Picking | |
| Delivering | |
| Selling | |
| Ripening | |
| Inventory of artificially-matured mangoes before ripening | |
| Inventory of artificially-matured mangoes with 2 days remaining shelf life | |
| Inventory of artificially-matured mangoes with 1 day remaining shelf life | |

**Table 5.** The optimal harvest decision in June (unit: tons)

| Varieties | 6/1-6/4 | 6/5 | 6/6 | 6/7 | 6/8 | 6/9 | 6/10 | 6/11 | 6/12 | 6/13 | 6/14 | 6/15 | 6/16 | 6/17 | 6/18 | 6/19 | 6/20 | 6/21 | 6/22 | 6/23 | 6/24 | 6/25 | 6/26 | 6/27 | 6/28 | 6/29 | 6/30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mangifera indica Linn | | [11,13] | [25,27] | [25,27] | [9,24] | [23,26] | [22,24] | [21,26] | [19,22] | [18,19] | [16,17] | [14,17] | [5,11] | | | [9,10] | | | [6,7] | [5,7] | [4,5] | | | | | | |
| | | | [11,13] | [25,27] | [25,27] | [24,29] | [23,26] | [22,24] | [20,26] | [19,22] | [18,19] | [16,17] | [14,17] | [5,11] | | | [9,10] | | | [6,7] | [5,6] | [0,4] | | | | | |
| | | | [11,13] | [25,27] | [25,27] | [24,29] | [23,26] | [22,24] | [20,26] | [19,22] | [18,19] | [16,17] | [14,17] | [5,11] | | | [9,10] | | | [6,7] | [5,6] | [0,4] | | | | | |
| Irwin | | | | | | | | | | | | | | | | | | | | [23,24] | [24,25] | [30] | [30] | [30] | [30] | [30] | [30] |
| | | | | | | | | | | | | | | | | | | | | | [23,24] | [24,25] | [30] | [30] | [30] | [30] | [30] |
| | | | | | | | | | | | | | | | | | | | | | [23,24] | [24,25] | [30] | [30] | [30] | [30] | [30] |
| Jin-Hwang | | | | | | | | | | | | | [17,21] | [28,30] | [3,5] | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | [54,56] | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | [54,56] | | | | | | | | |
| | | | | | | | | | | | | | | | [54,56] | | | | | | | | | | | | |
| | | | | | | | | | | | | | [18,21] | [48,51] | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | [53,56] | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Yu-Wen | | [16,19] | | | | | | | | | | | | | [24,26] | | | | | | | | | | | | |
| | | | | | | [16,19] | | | | | | | | | | | | | [24,26] | | | | | | | | |
| | | | | | | [16,19] | | | | | | | | | | | | | | | [24,26] | | | | | | |
| | | [16,19] | | | | | | | | | | | | | [24,26] | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | [16,19] | | | | | | | | | | | | | [24,26] | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Sensation | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Keitt | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

In **Table 7**, the Mangifera indica Linn is the earliest harvested mango, and its season starts at the beginning of June. All mangoes are delivered and sold on the next day after they are harvest. Another naturally matured mango, Irwin, is harvested from the end of June. Similar to Mangifera indica Linn, they are sold one day after harvest. For artificially matured mangoes, Yu-Wen mangoes are picked in early June, and Jin Hwang mangoes are harvested from June 18th to June 21st. Both mango types are ripened on the harvest day. Also, all mangoes are sold when the ripening process complete.

## 5.    Discussion and Conclusions

This study integrated forecasting and mathematical programming models for making mango harvest decisions in Taiwan. We addressed important concerns in the agriculture industry where demand and supply are highly uncertain. A time-series model was developed to forecast the demand and supply of different mango types using actual trading data in Taiwan. Also, a stochastic programming model was developed to determine the optimal picking, ripening, stocking, shipping, and selling decisions. The SAA was implemented for solving the stochastic program, and the computational result showed that the algorithm could obtain a quality solution in time. We conducted a case study to assess the value of stochastic solutions, where the stochastic programming model obtained more profitable solutions than the deterministic model used expected values. Finally, we organized the overall harvest plan according to the optimal solution to help farmers prepare resources in advance.

Our findings and suggestions are as follows. The workforce requirement to pick mangoes in June is higher than July and August. This is because the harvest time of naturally-matured mangoes is in June, and the harvest time cannot be changed. Thus, it is recommended to plant more mango varieties to alleviate peak labor requirements. The naturally-matured mangoes require less storage space because they are sold immediately after harvest. In contrast, more storage space should be planned for stocking artificially-matured mangoes to increase operational flexibility.

The limitation of this study is that we only consider operational decisions during a harvest season. Agriculture decisions in the early stage can have impacts. For example, mango yield quantity and timing are related to fruit bagging and pollination decisions made before the harvest season. A potential extension of our work is to include the above considerations to provide fruit farmers with more comprehensive planning.

# References

[1] Agriculture and Food Agency, Council of Agriculture, Executive Yuan. https://kmweb.coa.gov.tw/subject/ct.asp?xItem=177128&ctNode=2910&mp=110&kpi=0&hashid; 2017 [accessed March 18, 2017].

[2] Vizvári B, Lakner Z. A stochastic programming based analysis of the field use in a farm. Annals of Operations Research 2014; 219(1):231-242.

[3] Agriculture and Food Agency, Council of Agriculture, Executive Yuan. http://amis.afa.gov.tw/; 2017 [accessed April 19, 2017].

[4] Meuwissen M, Hardaker J, Huirne R, Dijkhuizen A. Sharing risks in agriculture; principles and empirical results. NJAS-Wageningen Journal of Life Sciences 2001; 49(4):343-356.

[5] Gupta A, Maranas CD. Managing demand uncertainty in supply chain planning. Computers & Chemical Engineering 2003; 27(8):1219-1227.

[6] Birge JR. Stochastic programming: computation and applications. INFORMS Journal on Computing 1997; 9:111-133.

[7] Birge JR, Louveaux F. Introduction to stochastic programming. 2nd ed. New York: Springer Science & Business Media; 2011.

[8] Ben-Tal A, El Ghaoui L, Nemirovski A. Robust optimization. New Jersey: Princeton University Press; 2009.

[9] Bohle C, Maturana S, Vera J. A robust optimization approach to wine grape harvesting scheduling. European Journal of Operational Research 2010; 200(1):245-252.

[10] Miller WA, Leung LC, Azhar TM, Sargent S. Fuzzy production planning model for fresh tomato packing. International Journal of Production Economics 1997; 53(3):227-238.

[11] Higgins AJ. Australian sugar mills optimize harvester rosters to improve production. Interfaces 2002; 32(3):15–25.

[12] Caixeta-Filho JV, van Swaay-Neto JM, Wagemaker ADP. Optimization of the production planning and trade of lily flowers at Jan de Wit Company. Interfaces 2002; 32(1):35-46.

[13] Caixeta-Filho JV. Orange harvesting scheduling management: a case study. Journal of the Operational Research Society 2006; 57:637-642.

[14] Vizvári B, Lakner Z, Csizmadia Z, Kovács G. A stochastic programming and simulation based analysis of the structure of production on the arable land. Annals of Operations Research 2011; 190(1):325-337.

[15] Guan Z, Philpott AB. A multistage stochastic programming model for the New Zealand dairy industry. International Journal of Production Economics 2011;

134(2):289-299.

[16] Ridier A, Chaib K, Roussy C. A dynamic stochastic programming model of crop rotation choice to test the adoption of long rotation under price and production risks. European Journal of Operational Research 2016; 252(1):270-279.

[17] Ahumada O, Villalobos JR. Operational model for planning the harvest and distribution of perishable agricultural products. International Journal of Production Economics 2011; 133(2):677-687.

[18] Ahumada O, Villalobos JR, Mason AN. Tactical planning of the production and distribution of fresh agricultural products under uncertainty. Agricultural Systems 2012; 112:17-26.

[19] Ferrer JC, Cawley AM, Maturana S, Toloza S, Vera J. An optimization approach for scheduling wine grape harvest operations. International Journal of Production Economics 2008; 112(2):985-999.

[20] Zhang W, Wilhelm WE. OR/MS decision support models for the specialty crops industry: a literature review. Annals of Operations Research 2011; 190(1):131-148.

[21] Lodree Jr EJ, Uzochukwu BM. Production planning for a deteriorating item with stochastic demand and consumer choice. International Journal of Production Economics 2008; 116(2):219-232.

[22] Noparumpa T, Kazaz B, Webster S. Wine futures and advance selling under quality uncertainty. Manufacturing & Service Operations Management 2015; 17(3):411-426.

[23] Bass FM. A new product growth for model consumer durables. Management science 1969; 15(5):215-227.