Agricultural Planning of Annual Plants Under Demand, Maturation, Harvest, and Yield Risk

Barış Tan
College of Administrative Sciences and Economics, Koç University, Rumeli Feneri Yolu, Istanbul, Turkey, 34450, btan@ku.edu.tr

Nihan Çömden
College of Engineering, Koç University, Rumeli Feneri Yolu, Istanbul, Turkey, 34450, nihancomden@gmail.com

In this study we present a planning methodology for a firm whose objective is to match the random supply of annual premium fruits and vegetables from a number of contracted farms and the random demand from the retailers during the planning period. The supply uncertainty is due to the uncertainty of the maturation time, harvest time, and yield. The demand uncertainty is the uncertainty of weekly demand from the retailers. We provide a planning methodology to determine the farm areas and the seeding times for annual plants that survive for only one growing season in such a way that the expected total profit is maximized. Both the single period and the multi period cases are analyzed depending on the type of the plant. The performance of the solution methodology is evaluated by using numerical experiments. These experiments show that the proposed methodology matches random supply and random demand in a very effective way and improves the expected profit substantially compared to the planning approaches where the uncertainties are not taken into consideration.

Key words: OR in Agriculture, Supply Chain Management, Inventory, Risk Management

1. Introduction

Matching supply and demand is an important problem in all industries. Several factors including the need to supply food to an increasing world population, decreasing farm lands, increasing food prices, and emergence of the premium food segment accentuate the importance of the management of agro-food supply chains today. Despite the technological advances that improve the effectiveness of supply chains in many industries, due to the biological nature of the agricultural production process, agro-food supply chains are prone to factors that are not completely controllable by humans. For agricultural production, planning decisions are made months before the actual demand is realized, with almost no flexibility to increase the supply afterwards. The quantity produced is affected by

the environmental conditions like weather, amount of sunlight and properties of soil. The weather conditions affect the growth of the seed, thereby the length of maturation and the harvest period during which crops are gathered. Although these risks cannot be controlled completely, they must be taken into consideration during the planning phase to match supply and demand in the most effective way.

In this study, we consider a planning problem in a premium fresh produce agricultural supply chain where a firm supplies fresh products to retailers by using contract farming. The firm's objective is to maximize its total profit over the planning period by determining the farm areas to be contracted and also the seeding times at the contracted farms. According to the contract between the firm and a number of selected farms, these farmers are provided with proprietary seeds and then paid according to the total output obtained at the end of each period. The output from each farm is collected by the firm, packed, and distributed to retailers every week during the planning period. The excess output at the end of each week cannot be stored or sold by the firm in alternate channels to ensure freshness of the products. Accordingly excess output is left to the farmers for their use in order to avoid unnecessary distribution and packaging costs. The profit is the difference between the revenue generated from sales to the retailers and the total cost of farming including the payments to the farmers, distribution, and packaging costs.

This planning problem is complicated due to various random factors that affect both supply and demand. The supply uncertainty is due to the uncertainty of the maturation time, harvest time, and yield. The total supply for a period consists of the total production in all farms harvested in that period. The length of the harvest period of a farm is random and different for each farm. In addition, the time of the harvest season differs based on the location of the farms. The company needs to use different farms in different locations to provide supply to its customers during the whole planning period. However, while planning the production period, the uncertainties in starting and ending times of the harvest period should also be considered. For instance, frost or rain during the maturation period or the harvest period can destroy all crops. Likewise undesirable weather conditions can draw out the growing season and cause the harvest season to start late. Furthermore,

all through the harvest the production quantity is random due to the random yield that highly depends on the weather and farm characteristics. The weekly demand from the retailers is also variable. These effects would cause the company to have no supply or less supply than desired and consequently the demand can be lost. The effective planning method presented in this study incorporates these risks and increase the firm's profits substantially compared to the planning approaches that do not take these risks into account.

The remainder of the paper is organized as follows: In the following section, the related literature is reviewed. In Section 3, the detailed description of the problem is given and the stochastic model for production planning of multiple farms under yield, demand and harvest uncertainty is presented. Before the agricultural planning problem is analyzed, a general solution to the newsvendor problem with normally distributed supply and demand is given in Section 4. In Section 5, we present the exact analytical solution of the single farm-single period problem for normally distributed yield and demand. We extend this result to the multi-farm single period case and develop accurate approximate solutions in Section 6. The multi-farm multi-period problem is discussed in Section 7. A case study in tomato farming is presented and the performance of the proposed approximate solutions are compared to the optimal and to a number of benchmark solutions in Section 8. Finally, conclusions are presented in Section 9.

2. Literature Survey

Agricultural planning activities can be grouped according to the planning horizon. For the long-range planning, farms must be chosen, crops must be selected, and a planting schedule must be determined, e.g. (Itoh et al. 2003), (Romero 2000), (Maatman et al. 2002), (Schilizzi and Kingwell 1999), (Jones et al. 2003). Once these decisions are given, shorter-range operational and tactical planning decisions including harvesting decisions such as how to allocate transportation equipment, labor force, scheduling of packing and processing plants, or amount of harvest per period can improve the performance significantly (Widodo et al. 2006), (Ferrer et al. 2008). Planning problems can also be grouped according to the type of plant that is of interest. Planning problems for annual

plants that survive for only one growing season and also for perennial plants that survive for more than one growing season are studied in the literature. In her thesis, Çömden (2009) presents the agricultural planning problem for annual and perennial plants with demand, maturation, harvest, and yield uncertainty. The reader is referred to the comprehensive review of Ahumada and Villalobos (2009) for the pertinent literature in the area of production planning problems in agricultural supply chains.

Our study is related to the agricultural planning models where both supply and demand risks are incorporated into the problem. Kazaz (2004) considers a production planning problem of a company that produces olive oil in the case of random yield and demand. Allen and Schuster (2004) develop a mathematical model to control the harvest risk in grape farming by scheduling of the harvest.

Our problem is a long-range planning problem for annual plants where seeding times and farm areas are determined. Merrill (2007) presents a similar contract farm selection and area determination problem for premium tomatoes. Once the farm areas are determined and the seeds are planted according to the optimal seeding times, harvest decisions can be optimized by using tactical planning approaches such as the methods presented in (Ahumada and Villalobos 2011a) and(Ahumada and Villalobos 2011b. These two papers also consider variability in harvest and maturation times and provide a planning method that manages the tradeoff between the added labor and distribution costs and the freshness at delivery of the products.

In its essence, the problem considered in this study is an inventory planning problem with random supply and demand. Yano and Lee (1995) present an extensive literature review of lot sizing models with random yield. Gerchak et al. (1988) present periodic review production models with random yield and demand.

In the solution of the single-period problem, we consider the newsvendor problem with random yield. Anupindi and Akella (1993) study the strategy of supplying from two suppliers. Dada et al. (2007) consider the procurement problem of a newsvendor when the suppliers are unreliable. Yang

et al. (2007) examine a supplier selection problem, where decision maker orders from a set of suppliers with different yields while facing a random demand.

Rekik et al. (2007) study the newsvendor problem with unreliable supply and investigate the optimal order quantities for different cases. We generalize the solution of Rekik et al. (2007) for the cases where both the mean and the standard deviation are written as a function of a variable.

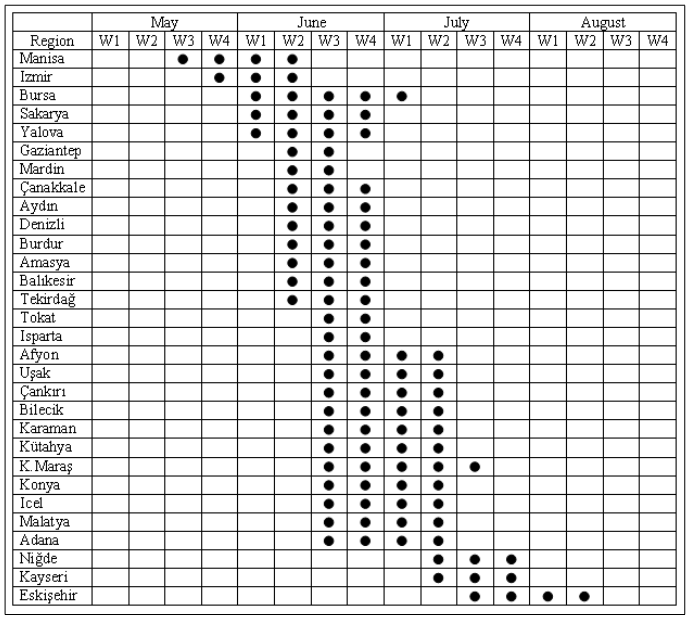
There are two main contributions of this study. First, an analytical planning model for agricultural production with harvest, maturation, yield and demand uncertainty is presented and the analytic solutions of single-farm single-period, multiple farm-single period, and multiple farm-multiperiod problems are developed. Second, our work contributes to random yield, random demand problems by providing a closed form solution of the newsvendor problem when the demand and the supply are normal random variables and the mean and the standard deviation of the supply are functions of a decision variable.

3. Agricultural Planning Problem

We consider a firm that evaluates N different farms located in different regions. We use a discrete model where each period has the same length, typically a week. The planning horizon consists of $\mathcal{T}$ periods. Figure 1 depicts the weeks in a year where cherry farms located in different regions of Turkey are available for harvest. A company that supplies agricultural products to retailers throughout the year must use different farms located in different regions. In addition, since agricultural processes are random by nature, there is always a risk in the availability of the crop in each period for each farm. The production quantity at each farm, $q_i(t)$ is random and depends on the planted area and the crop yield in that farm. The total production quantity, $Q(t)$, is the total amount of crop supplied from all farms in period t, $Q(t) = \sum_{i=1}^{N} q_i(t)$.

The area to be seeded in each contracted farm i is a_i. The period where farm i is seeded is referred as the seeding time and denoted with τ_i, $i = 1, ..., N$. In the planning problem, a_i and τ_i, $i = 1, ..., N$ are the decision variables.

Yield denoted with $Y_i(t)$ is defined as the amount of output gained from one acre of seeded farm in period t. When the harvest starts at farm i, an output of $a_i Y_i(t)$ is obtained in period t. We



	May				June				July				August			
Region	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4
Manisa			●	●	●	●										
Izmir				●	●	●										
Bursa					●	●	●	●	●							
Sakarya					●	●	●	●								
Yalova					●	●	●	●								
Gaziantep						●	●									
Mardin						●	●									
Çanakkale						●	●	●								
Aydın						●	●	●								
Denizli						●	●	●								
Burdur						●	●	●								
Amasya						●	●	●								
Balıkesir						●	●	●								
Tekirdağ						●	●	●								
Tokat							●	●								
Isparta							●	●								
Afyon							●	●	●	●						
Uşak							●	●	●	●						
Çankırı							●	●	●	●						
Bilecik							●	●	●	●						
Karaman							●	●	●	●						
Kütahya							●	●	●	●						
K.Maraş							●	●	●	●	●					
Konya							●	●	●	●						
Icel							●	●	●	●						
Malatya							●	●	●	●						
Adana							●	●	●	●						
Niğde										●	●	●				
Kayseri										●	●	●				
Eskişehir											●	●	●	●		

Figure 1 The harvest availabilities of cherry in Turkey

model $Y_i(t)$ as a normal random variable with mean $E[Y_i(t)]$, variance $Var[Y_i(t)]$, and the standard deviation $\sigma[Y_i(t)]$.

The demand in period t is a random variable denoted with $D(t)$. The mean, the variance, and the standard deviation of the demand are $E[D(t)]$, $Var[D(t)]$, and $\sigma[D(t)]$ respectively.

3.1. Optimization Problem

The objective of the firm is to maximize the total expected profit of a single product that is produced by multiple farms of an agricultural supply chain system over a planning period of $\mathcal{T}$ periods.

The firm generates a revenue of $r(t)$ from sales of one unit of output to the retailer in period t. According to the contract, each contracted farm is paid for the whole production. The total production, distribution, and packaging cost of supplying one unit of production from each farm in period t is $c(t)$. If the total supply is greater than the demand in a given period, the firm leaves the excess production at the farms in order not to incur unnecessary distribution and packaging costs. The distribution and packaging cost per unit of output is denoted with $s(t)$. We assume $r(t) > c(t) > s(t)$.

The optimization problem is determining the farm area a_i and the seeding time τ_i, $i = 1, ..., N$

for each farm such that the total expected profit $E[\pi(\mathbf{a},\tau)]$ is maximized:

$$\begin{aligned}\underset{\mathbf{a},\tau}{\text{Max}} \quad E[\pi(\mathbf{a},\tau)] &= \sum_{t=1}^{\mathcal{T}} r(t)E[\min(Q(t),D(t))] + s(t)E[(Q(t)-D(t))^{+}] - c(t)E[Q(t)] \\ &= \sum_{t=1}^{\mathcal{T}} (r(t)-s(t))E[\min(Q(t),D(t))] - (c(t)-s(t))E[Q(t)] \qquad (1)\end{aligned}$$

where $\mathbf{a}=(a_1,a_2,...,a_N)$, and $\tau=(\tau_1,\tau_2,...,\tau_N)$, and $(Q(t)-D(t))^{+}=\max\{Q(t)-D(t),0\}$.

3.2. Supply Model

We use a product form expression for the supply from farm i at time t:

$$q_i(t) = I_{i,\tau_i}(t)Y_i(t)a_i$$

where $I_{i,\tau_i}(t)$ is an indicator variable that is 1 if the maturation time is completed and the farm is in its harvest period and farm i is available for harvesting, i.e., the farm is in its harvest period in period t when the seeds are planted in period τ_i. By using available information such as the cumulative output per plant per day, the parameters of the distributions of $I_{i,\tau_i}(t)$ and $Y_i(t)$ can be estimated.

The lengths of the maturation and the harvest period are the two important factors for the output availability. These times are uncertain and highly dependent on weather. Using only the expected values in the planning can yield considerable deviation from the actual realization. The farm availability probability $p_{i,\tau_i}(t)$ is defined as the probability that the maturation duration of the farm is completed and an output from farm i is available for harvest in period t given that the seeds are planted in period τ_i. Since $p_{i,\tau_i}(t)$ is defined as the probability that the output from farm i will be available in period t given that the seeds are planted in period τ_i, $E[I_{i,\tau_i}(t)]=p_{i,\tau_i}(t)$.

The form and definition of the farm availability probability $p_{i,\tau_i}(t)$ allow us to use this parameter directly either by using past data or by incorporating expert opinion. Specifying $p_{i,\tau_i}(t)$ directly makes it easier to capture harvest risk dependencies among different farms possibly located in the same region.

Alternatively, the farm availability probability can be derived by using the distributions that represent the random maturation and harvest times. The length of the maturation and the harvest period are random variables and they are denoted with M_{i,τ_i} and H_{i,τ_i} respectively.

Figure 2 shows the timeline of the events in the model. We assume that the farms are seeded at the beginning of a period; the demand is realized at the beginning of each period; and the crops are gathered at the end of each period. Since no inventory can be kept due to the perishability of the product, the crops harvested in one period can only be used to satisfy the demand of the same period.

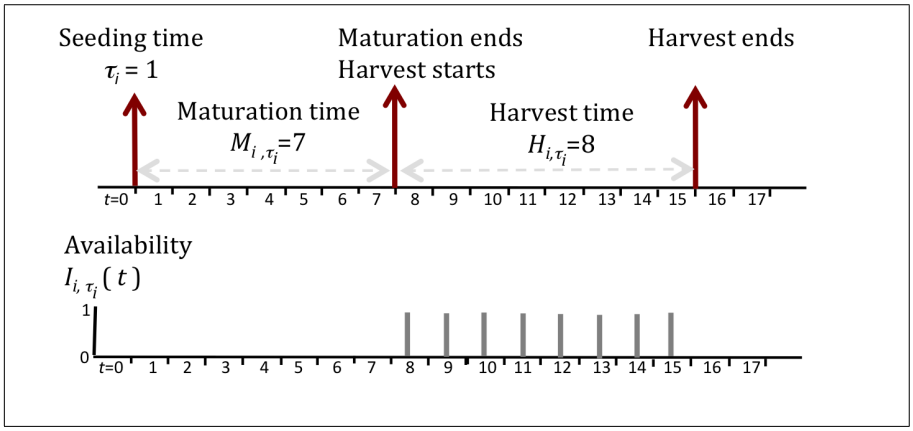


Figure 2 The relation between the seeding time, the maturation and the harvest period

Accordingly,the indicator variable for the availability of farm i at time t $I_{i,\tau_i}(t)$ can be expressed in terms of τ_i, M_{i,τ_i} and H_{i,τ_i} as

$$I_{i,\tau_i}(t) = \begin{cases} 1 & \tau_i + M_{i,\tau_i} \leq t < \tau_i + M_{i,\tau_i} + H_{i,\tau_i} \\ 0 & otherwise \end{cases}. \tag{2}$$

If the probability density functions of M_{i,τ_i} and H_{i,τ_i} are estimated from the data, the above function can be used to determine $p_{i,\tau_i}(t) = E[I_{i,\tau_i}(t)]$.

We assume that the firm has sufficient resources so that the harvest time is independent of the farm area. Since we are studying a problem in a supply chain for premium vegetables and fruits, firms make the necessary investments to complete the harvest without any resource conflicts. Note that once the long-range decisions are given in the best way to match supply and

demand, resources that are necessary for the harvest can be managed effectively by using the tactical planning approaches (Ahumada and Villalobos 2011a).

The total supply from a farm during the harvest period is dependent on the yield and also on the length of the harvest period. Therefore the moments of the distributions of the yield and the harvest length are interrelated to each other to give the observed moments of the total yield during the harvest period. Since $I_{i,\tau_i}(t)$ and $Y_i(t)$ are random variables, $Q(t)$ is also a random variable. The mean and variance of $Q(t)$ can be written as

$$E[Q(t)] = \sum_{i=1}^{N} E[q_i(t)] = \sum_{i=1}^{N} p_{i,\tau_i}(t) E[Y_i(t)] a_i \tag{3}$$

and

$$Var[Q(t)] = \sum_{i=1}^{N} \left(p_{i,\tau_i}(t) E[Y_i(t)^2] - p_{i,\tau_i}^2(t) E^2[Y_i(t)] \right) a_i^2. \tag{4}$$

4. Newsvendor Problem with Random Supply

Before we present the solution methodologies for the problem given in Equation (1), we first focus on the newsvendor problem with normal supply and normal demand. We use this model to get insights and to derive general results that will be used in the solution of the agricultural planning problem.

In the newsvendor problem with random supply, the decision maker's order decision yields a supply quantity that is random with a distribution that depends on the decision. Let the normally distributed random variables for the supply and demand be denoted with Q and D respectively. We assume that the mean and the standard deviation of the supply quantity Q are functions of a single decision variable x, $E[Q] = f(x)$ and $\sigma[Q] = g(x)$ respectively. For example, in the newsvendor problem with a multiplicative random yield, when an order with a quantity of x is given, the supply quantity obtained is $x\epsilon$ where ϵ is a random variable with mean μ_ϵ and standard deviation σ_ϵ. Accordingly, the supply quantity decision x yields a supply that has a mean of $x\mu_\epsilon$ and a standard deviation of $x\sigma_\epsilon$. Therefore, we can set $f(x) = x\mu_\epsilon$ and $g(x) = x\sigma_\epsilon$ to analyze this case. In the agricultural planning problem, the farm area decision affects both the mean and also the standard deviation of the supply according to Equations (3) and (4).

After dropping the subscript i and t, the expected profit for a single-period newsvendor problem with random supply can be written as

$$E\left[\pi(x)\right] = (r-s)E[\min\left(Q,D\right)] - (c-s)E[Q] \tag{5}$$

where $E[Q] = f(x)$ and $\sigma[Q] = g(x)$. In our analysis, we utilize a result for the expected value of the minimum of two correlated normal random variables (Cain 1994). Accordingly, the expected value of the minimum of two correlated normal random variables can be written in closed form as

$$E[\min\left(Q,D\right)] = E[Q]\Phi\left(\frac{E[D]-E[Q]}{\theta}\right) + E[D]\Phi\left(\frac{E[Q]-E[D]}{\theta}\right) - \theta\phi\left(\frac{E[D]-E[Q]}{\theta}\right) \tag{6}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $\phi(\cdot)$ is the standard normal probability density function, $\theta = \left(Var[Q] - 2\rho\sqrt{Var[Q]Var[D]} + Var[D]\right)^{1/2}$ and ρ is the correlation coefficient between Q and D.

We assume that uncertainties affecting the supply and demand are independent. Accordingly, by using Equation (6) with $\rho = 0$, the expected profit for the newsvendor problem with normal supply and normal demand can be written as a function of the decision variable x as

$$\begin{aligned} E[\pi(x)] = {} & (r-s)\left[f(x)\Phi\left(\frac{E[D]-f(x)}{h(x)}\right) + E[D]\Phi\left(\frac{f(x)-E[D]}{h(x)}\right) - h(x)\phi\left(\frac{E[D]-f(x)}{h(x)}\right)\right] \\ & - (c-s)\,f(x) \end{aligned} \tag{7}$$

where $h(x) = \sqrt{g^2(x) + Var[D]}$.

By using the above closed-form expression for the expected total profit, the decision variables that maximize the expected total profit can be determined. The following theorem shows that the optimal x can be determined by solving a nonlinear equation.

THEOREM 1. *For the newsvendor problem with normally distributed supply and demand, if $E[Q] = f(x)$ is linear and $\sigma[Q] = g(x)$ is convex, the expected profit function is guaranteed to be a concave function and the optimal x that maximizes the expected profit satisfies the following equality*

$$f(x) = E[D] - h(x)\Phi^{-1}\left[\frac{c-s}{r-s} + g(x)\frac{\partial g(x)}{\partial x}h(x)^{-1}\phi\left(\frac{E[D]-f(x)}{h(x)}\right)\left(\frac{\partial f(x)}{\partial x}\right)^{-1}\right] \tag{8}$$

The proof is given in the Appendix. By using the above result, we analyzed the specialized cases of the newsvendor problem for additive, multiplicative, and additive-multiplicative random yield cases and showed that the results were the same as the optimal quantities presented in the analysis of Rekik et al. (2007). The main difference between the analysis of Rekik et al. (2007) and our result is that we use a more general approach to model supply uncertainty where the mean and variance of the supply quantity are functions of the decision variable. We use Theorem 1 to state the optimal solution of the single-farm, single-period problem in the next section.

5. Single-Farm, Single-Period Problem

In this section, we consider the single period case of the problem given in Equation (1) for a single farm. The single-period problem is of interest for planning problems for seasonal products and provides structural results that can be used for the general case. In this case, the demand exists only for a single period t^*. The planning problem of annual plants is finding the optimal a_1 and τ_1 to maximize the expected profit in period t^* that is given as

$$E[\pi(a_1,\tau_1)] = r(t^*)E[\min(Q(t^*),D(t^*))] + s(t^*)E[(Q(t^*)-D(t^*))^+] - c(t^*)E[Q(t^*)]. \quad (9)$$

This problem is a newsvendor problem with random demand and random supply. However, since the distribution of the supply quantity is not exactly normal in our planning problem, the single-farm single-period problem differs from the newsvendor problem with random supply problem described in Section 4.

The supply quantity depends on the availability of harvest in the related period which is a binary random variable and also on the farm yield which is a normal random variable. In period t the output quantity from farm i is equal to $Y_i(t)a_i$ with probability $p_{i,\tau_i}(t)$ and it is 0 with probability $1-p_{i,\tau_i}(t)$. For the case when there is only one farm, $\min(Q(t),D(t))$ can be expressed as

$$\min(Q(t),D(t)) = \begin{cases} 0 & \text{with probability } 1-p_{1,\tau_1}(t) \\ \min(Q'(t),D(t)) & \text{with probability } p_{1,\tau_1}(t) \end{cases}$$

where $Q'(t) = a_1 Y_1(t)$. Since $Y_1(t)$ is normal, Equation (6) can be rewritten as

$$E[\min(Q(t),D(t))] = p_{1,\tau_1}(t)\left(E[Q'(t)]\,\Phi\left(\frac{E[D(t)]-E[Q'(t)]}{\theta'(t)}\right) + E[D(t)]\,\Phi\left(\frac{E[Q'(t)]-E[D(t)]}{\theta'(t)}\right)\right.$$

$$-\theta(t)\phi\left(\frac{E\left[D(t)\right]-E\left[Q'(t)\right]}{\theta'(t)}\right)\Bigg)$$

where $E\left[Q'(t)\right]=E\left[Y_1(t)\right]a_1$, $Var\left[Q'(t)\right]=Var\left[Y_1(t)\right]a_1^2$ and $\theta'(t)=\sqrt{Var\left[Y_1(t)\right]a_1^2+Var\left[D(t)\right]}$.

Then the expected profit for period t^* given in Equation (9), can be written as

$$\begin{aligned}E\left[\pi\left(a_1,\tau_1\right)\right]=&\left(r(t^*)-s(t^*)\right)p_{1,\tau_1}(t^*)\Bigg(E\left[Q'(t^*)\right]\Phi\left(\frac{E\left[D(t^*)\right]-E\left[Q'(t^*)\right]}{\theta'(t^*)}\right)\\&+E\left[D(t^*)\right]\Phi\left(\frac{E\left[Q'(t^*)\right]-E\left[D(t^*)\right]}{\theta'(t^*)}\right)-\theta'(t^*)\phi\left(\frac{E\left[D(t^*)\right]-E\left[Q'(t^*)\right]}{\theta'(t^*)}\right)\Bigg)\\&-\left(c(t^*)-s(t^*)\right)p_{1,\tau_1}(t^*)E[Q'(t^*)].\end{aligned}\tag{10}$$

There are two decision variables in the problem: a_1 and τ_1. The optimal farm area is determined by using the general result given in Theorem 1 for the newsvendor problem with normally distributed supply and demand.

THEOREM 2. *The optimal farm area that maximizes the expected profit for a given seeding time in the single-farm single-period problem satisfies the following equality*

$$a_1=\frac{E[D(t^*)]}{E\left[Y_1(t^*)\right]}-\frac{\theta'(t^*)}{E\left[Y_1(t^*)\right]}\Phi^{-1}\left(\frac{\left(c(t^*)-s(t^*)\right)}{\left(r(t^*)-s(t^*)\right)}+\frac{a_1Var\left[Y_1(t^*)\right]}{E\left[Y_1(t^*)\right]\theta'(t^*)}\phi\left(\frac{E[D(t^*)]-E[Y_1(t^*)a_1]}{\theta'(t^*)}\right)\right)\tag{11}$$

where $\theta'(t^*)=\sqrt{Var\left[Y_1(t^*)\right]a_1^2+Var\left[D(t^*)\right]}$.

The proof is given in the Appendix.

From the equation above, the optimal decision for the farm area can be found for a given seeding time decision. Next we give the result for the optimal seeding time for the problem of annual plants.

THEOREM 3. *The optimal seeding time is the time that yields the highest farm availability probability at the desired period* t^*.

The proof is given in the Appendix.

The most important managerial insight from Theorems 2 and 3 is that the seeding area decision can be separated from the seeding time decision. More specifically, the farm availability probability is dependent on the maturation time harvest time uncertainties. Therefore the maturation and harvest time uncertainty can be mitigated by determining the seeding time while the yield and demand uncertainty can be mitigated by determining the farm area.

6. Multi-Farm, Single-Period Problem

For the planning problem for the multi-farm single-period setting, there are multiple farms available and before the planning period starts, the decision maker needs to determine the seeding areas and the seeding times for those farms to maximize the profit in period t^*.

Figure 3 shows a sample distribution of $q_i(t)$ and $Q(t)$ obtained by simulation for a particular case with 3 farms over 10 periods.

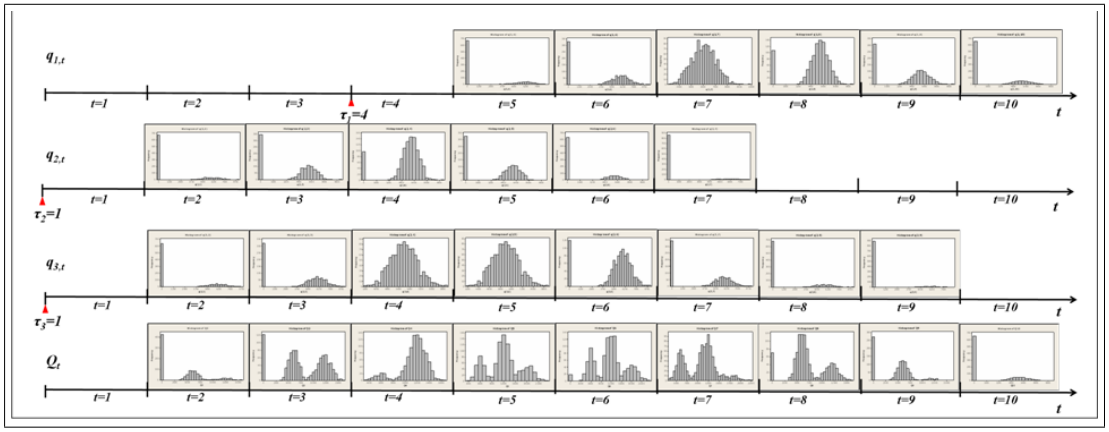


Figure 3 The histogram of $q_i(t)$s and $Q(t)$

The expected profit for period t^* is written as

$$E[\pi(\mathbf{a},\tau)] = (r(t^*) - s(t^*))E[\min(Q(t^*), D(t^*))] - (c(t^*) - s(t^*))E[Q(t^*)]. \tag{12}$$

6.1. Exact Solution of the Multi-Farm, Single-Period Problem

By using the farm availability probabilities, the expected total output from N farms can be written as

$$E[Q(t^*)] = \sum_{\alpha_1=0}^{1} \cdots \sum_{\alpha_N=0}^{1} \left(\prod_{i=1}^{N} p_{i,\tau_i}^{\alpha_i}(t^*) \left(1 - p_{i,\tau_i}(t^*)\right)^{1-\alpha_i} \right) E\left[\sum_{i=1}^{N} q_i'(t^*)\alpha_i \right].$$

where $q_i(t)' = Y_i(t)a_i$.

Then the the expected profit from N farms for a single period can be expressed as

$$E[\pi(\mathbf{a},\tau)] = \sum_{\alpha_1=0}^{1} \cdots \sum_{\alpha_N=0}^{1} \left(\prod_{i=1}^{N} p_{i,\tau_i}^{\alpha_i}(t^*) \left(1 - p_{i,\tau_i}(t^*)\right)^{1-\alpha_i} \right) \left((r(t^*) - s(t^*))E\left[\min\left(\sum_{i=1}^{N} q_i'(t^*)\alpha_i, D(t^*) \right) \right] - (c(t^*) - s(t^*))E\left[\sum_{i=1}^{N} q_i'(t^*)\alpha_i \right] \right) \tag{13}$$

Since $q_i'(t)$ is normally distributed, the expansion of the expectation of the minimum of two normal random variables given in Equation (6) yields a closed-form expression for Equation (13). We next show that this expected profit is a concave function of farm areas. As a result, maximization of Equation (13) yields the desired optimal farm areas.

THEOREM 4. *$E[\pi(\mathbf{a},\tau)]$ is a concave function of $\mathbf{a}$. For all N values, $N, \in R^+$, there exist a unique $\mathbf{a} = (a_1, ..., a_N)$, that maximizes the expected profit function, $E[\pi(\mathbf{a},\tau)]$.*

The proof is given in the Appendix.

In numerical experiments, it is observed that Theorem 3 also holds for the optimal seeding time for each farm in the multi-farm problem. Namely, the optimal seeding time for each farm in the multi-farm single-period case is the time that yields the highest farm availability probability for that farm at the desired period t^*. A formal proof of this result is left for future research.

6.2. Normal Approximation for The Solution of Multi-Farm, Single-Period Problem

For the general case, the number of computations to evaluate the profit function in Equation (13) increases exponentially as the number of farms increases. Since the total supply is the sum of the output from different farms, we expect that the distribution of the total supply approaches to the normal distribution as the number of farms increases. Accordingly, we approximate the supply as a normally distributed random variable in each period t.

Under the approximation that the total supply is normally distributed, expanding Equation (1) directly by using Equation (6) yields

$$\tilde{\pi}(\mathbf{a},\tau) = (r(t^*) - s(t^*))\left(E[Q(t^*)]\,\Phi\left(\frac{E[D(t^*)] - E[Q(t^*)]}{\theta(t^*)}\right) + E[D(t^*)]\,\Phi\left(\frac{E[Q(t^*)] - E[D(t^*)]}{\theta(t^*)}\right) - \theta(t^*)\phi\left(\frac{E[D(t^*)] - E[Q(t^*)]}{\theta(t^*)}\right)\right) - (c(t^*) - s(t^*))E[Q(t^*)] \quad (14)$$

where $\tilde{\pi}$ is the expected profit obtained when $Q(t)$ is approximated with a normal random variable with the mean and the variance provided in Equations (3) and (4).

Since this case is identical to the newsvendor problem with normal supply and demand when the seeding times are given, Theorem 1 yields the desired farm sizes directly. Using $E[Q(t)]$ and

$Var[Q(t)]$ in Equation (8) yields the farm area that maximizes the expected profit under the normality assumption of the total supply, $\tilde{a}_i^*$.

THEOREM 5. *For given seeding times, the optimal farm areas that maximize the expected profit for the multi-farm single-period under the normality assumption of the total supply satisfy the following set of equations for* $i = 1, .., .N$*:*

$$\tilde{a}_i^* = \frac{1}{p_{i,\tau_i}(t^*)E[Y_i(t^*)]}\Bigg[E[D(t^*)] - \theta(t^*)\Phi^{-1}\left(\frac{c(t^*) - s(t^*)}{r(t^*) - s(t^*)}\right. \\ \left. - \frac{(E[Y_i(t^*)^2] - p_{i,\tau_i}(t)E^2[Y_i(t^*)])a_i}{E[Y_i(t^*)]\theta(t^*)}\phi\left(\frac{E[D(t^*)] - E[Q(t^*)]}{\theta(t^*)}\right)\right)\Bigg]. \quad (15)$$

Since the result follows Theorem 1 directly, the proof is omitted.

We next give the result for the best seeding time that maximizes the expected profit under the normality assumption of the total supply.

THEOREM 6. *The seeding time that maximizes the expected profit under the normality assumption of the total supply is the time that yields the highest farm availability probability at the desired period* t^**.*

The proof is given in the Appendix.

Note that Theorem 3 and Theorem 6 suggest that both for the exact solution and for the approximated solution, the best seeding time decision is the one that gives the maximum harvest probability for the target period. As a result, the seeding time obtained with the normality assumption is optimal.

In order to evaluate the performance of the approximate solution of the multi-farm single-period case for annual plants, we analyzed 480 cases with different number of farms, expected yield, expected demand, production cost and harvest probability values. The parameters of scenarios are given in Table 1. In the experiments, all farms are assumed to have the same harvest, maturation, and yield characteristics. For each case, we obtained the optimal solution by determining the optimal farm area by maximizing the expected profit given in Equation (13) for all possible seeding times. Figure 4 shows the percentage difference between the optimal solution and the approximated

Parameters	Values
N :	$1, ..., 8$
$r(t^*)$:	5
$c(t^*)$:	$2, 3$
$s(t^*)$:	0.5
$E[D(t^*)]$:	2000
$\sigma[D(t^*)]$:	$200, 400, 600$
$max(p_{i,\tau_i}(t^*))$:	$1, 0.9, 0.8, 0.7, 0.6$
$E[Y_i(t^*)]$:	6000
$\sigma[Y_i(t^*)]$:	$600, 1200$

Table 1 **Parameters of Scenarios**

solution of the multi-farm single-period problem for different N values. Figure 4 shows that the difference between the optimal solution and the approximate solution is 2.3% when there is only one farm and it decreases to 0.001% rapidly as the number of farms increases to 8. As a result, this approximation method can be used effectively.

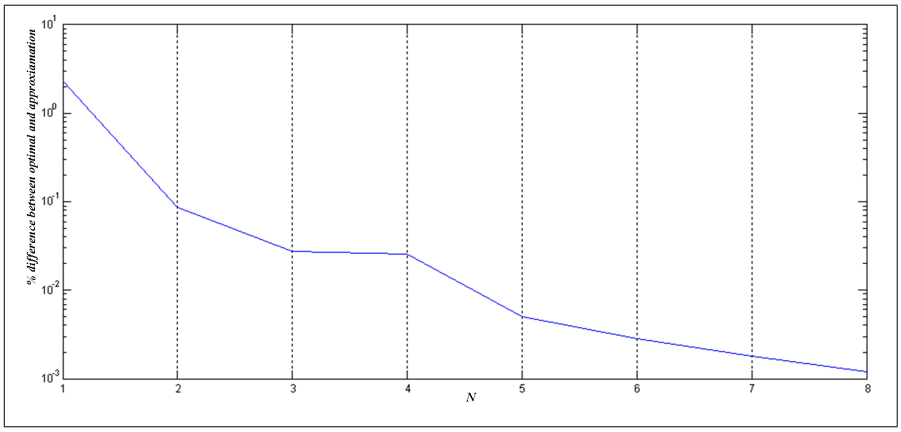


Figure 4 **The relationship between the number of farms and the percentage error between the optimal solution and the approximation for the Multi-farm Single-period problem**

7. Multi-Farm Multi-Period Problem

We now consider the planning problem when the planning period consists of $\mathcal{T}$ periods and the number of available farms is N. In the beginning of the planning period, the decision maker needs to decide the farm areas to be seeded and the seeding times to maximize the expected total profit during the planning period.

7.1. Exact Solution of the Multi-Farm, Multi-Period Problem

The planning problem of annual plants where the decision variables are farm areas and seeding times can be expressed as a nonlinear assignment problem given below:

$$\underset{\mathbf{a},\tau}{\text{Max}} \quad E\left[\pi\left(\mathbf{a},\tau\right)\right] \tag{16}$$

subject to

$$\begin{aligned}
\tau_i &= \sum_{t=1}^{\mathcal{T}} tW_i(t), && i=1,...,N \\
\sum_{t=1}^{\mathcal{T}} W_i(t) &\leq 1, && i=1,...,N \\
W_i(t) &\in \{0,1\}, && i=1,...,N; t=1,...\mathcal{T}
\end{aligned} \tag{17}$$

where $E[\pi\left(\mathbf{a},\tau\right)]$ is the total expected profit in $\mathcal{T}$ periods. In the above formulation, the seeding time τ_i is expressed as an assignment decision where a binary decision variable $W_i(t)$ determine in which period from $t=1$ to $\mathcal{T}$ farm i is seeded.

The total profit of all periods is equal to the sum of profits in each period. Since it is assumed that no inventory is carried from one period to another in order to preserve the freshness of the premium products, the expected total profit for the multi-period case is the sum of the expected profits for a single period given in Equation (13) for all the time periods:

$$\begin{aligned}
E[\pi(\mathbf{a},\tau)] = &\sum_{t=1}^{\mathcal{T}}(r(t)-s(t)) \sum_{\alpha_i=0}^{1} ... \sum_{\alpha_N=0}^{1} \prod_{i=1}^{N} p_{i,\tau_i}^{\alpha_i}(t)(1-p_{i,\tau_i}^{1-\alpha_i}(t)) E\left[\min\left(\sum_{i=1}^{N} q_i'(t)\alpha_i, D(t)\right)\right] \\
&- (c(t)-s(t))E[Q(t)].
\end{aligned} \tag{18}$$

This expected profit is a concave function of farm areas as shown in Theorem 7.

THEOREM 7. *$E[\pi(\mathbf{a},\tau)]$ is a concave function of $\mathbf{a}$ and there exists a unique $\mathbf{a}=(a_1,...,a_N)$, that maximizes the expected profit function, $E[\pi(\mathbf{a},\tau)]$.*

Proof. Theorem 4 proves that $E[\pi(\mathbf{a},\tau)]$ is a concave function of $\mathbf{a}$ when $\mathcal{T}=1$. Since the sum of concave functions is concave, the expected profit for the multi-period multi-farm case given in Equation (18) is concave. □

As a result, maximization of Equation (18) yields the desired optimal farm areas.

7.2. Normal Approximation for the Solution of Multi-Farm, Multi-Period Problem

Since the planning problem is a mixed-integer nonlinear optimization problem where the the objective function has an evaluation complexity of $O(2^n)$, direct solution is not computationally feasible. We propose using an approximate solution to this problem where the normal approximation for the total supply is used and an iterative solution procedure is employed to determine the farm areas and seeding times.

Accordingly, we approximate the supply quantity in each period t as a normally distributed random variable with mean $E[Q(t)]$ and variance $Var[Q(t)]$ given in Equations (3) and (4). By using Equation (7), the expected profit can be written as;

$$\tilde{\pi}(\mathbf{a},\tau) = \sum_{t=1}^{\mathcal{T}} (r(t)-s(t)) \left(E[Q(t)]\Phi\left(\frac{E[D(t)]-E[Q(t)]}{\theta(t)}\right) + E[D(t)]\Phi\left(\frac{E[D(t)]-E[Q(t)]}{\theta(t)}\right) - \theta\phi\left(\frac{E[D(t)]-E[Q(t)]}{\theta(t)}\right)\right) - (c(t)-s(t))E[Q(t)]. \tag{19}$$

Using the normality assumption for total supply leads to a nonlinear integer model where the objective function in Equation (16) is replaced with the concave objective function given in Equation(19) and solved with the assignment constraints given in Equation (17) .

In order to improve the computational efficiency compared to solving a nonlinear integer optimization problem, we propose a two-stage approach where in the first stage the farm areas are determined for given seeding time. When the farm areas are given, the model given in Equations (16)- (17) is a linear assignment problem. By solving this assignment problem in the second stage, the best seeding times are determined for the given seeding areas. These two problems are solved iteratively to obtain an approximate solution to the multi-farm multi-period planning problem for annual plants.

7.3. Mean-value Solution of the Multi-farm Multi-period Problem

In the numerical experiments, we compare the solution of the stochastic planning problem with the solution obtained by using a mean-value approach. When the planning problem is modelled by using only the expected values of the random variables for yield, demand, maturation and harvest

length, the indicator variable $I_{i,\tau_i}(t)$ becomes a definite parameter, $I'_{i,\tau_i}(t)$, and it can be written as

$$I'_{i,\tau_i}(t) = \begin{cases} 1 & \tau_i + E[M_{i,\tau_i}] \leq t < \tau_i + E[M_{i,\tau_i}] + E[H_{i,\tau_i}] \\ 0 & otherwise \end{cases} \tag{20}$$

Then, the optimization problem for the multi-farm, multi-period problem when the random variables are replaced with their expected values is given as

$$\underset{\mathbf{a},\tau}{\text{Max}} \quad \sum_{t=1}^{\mathcal{T}} r(t)\min(\bar{Q}(t), E[D(t)]) + s(t)(\bar{Q}(t) - E[D(t)])^{+} - c(t)\bar{Q}(t) \tag{21}$$

subject to

$$\begin{aligned} \bar{Q}(t) = \sum_{i=1}^{N} I'_{i,\tau_i}(t) E\left[Y_i(t)\right] a_i, \qquad & t = 1, ...\mathcal{T} \\ \tau_i = \sum_{t=1}^{\mathcal{T}} t W_i(t), \qquad & i = 1, ..., N \\ \sum_{t=1}^{\mathcal{T}} W_i(t) \leq 1, \qquad & i = 1, ..., N \\ W_i(t) \in \{0, 1\}, \quad & i = 1, ..., N; t = 1, ...\mathcal{T} \end{aligned} \tag{22}$$

The above problem is a nonlinear maximization problem. However it can be rewritten as a linear programming problem by defining $\bar{Q}(t) - E\left[D(t)\right] = \Delta_t^+ - \Delta_t^-$ where $\Delta_t^+ \geq 0, \Delta_t^- \geq 0$.

8. A Case Study and Numerical Results

In this section, we first discuss the case of tomato farming and focus on data availability and input parameter selection. Then we evaluate the performance of the proposed solution methodology by comparing the results with the results obtained by using an approach that does not take risks into account.

8.1. A Case Study in Tomato Farming

In order to explain the problem, we discuss the case of a firm that buys premium tomato seeds from a seed producer and then contracts a number of farms in different geographical regions to plant these seeds and produce premium fruits and vegetables. Once tomatoes are harvested, the output is transported from all contracted farms to the packaging and processing facilities by the

firm and then distributed to retailers. A similar agribusiness supply chain for tomato farming is also described by Merrill (2007). Tan (2011) presents the case of Alara Agribusiness in cherry farming.

The prices of tomato seeds vary from \$0.01 per seed for commodity tomatoes to \$4 per seed for specialty tomatoes. For example the seeds of Summer Sun Yellow that is a specialty premium tomato are sold by a Biotech firm at a price of \$350,000 per kg or \$0.9 per seed. Each seed can yield an output of 20 kg that can be sold at a retail price of \$26 per kg. An intermediary firm buys these seeds and distributes them to a number of farms with a contract that specifies the farm area to be used to plant and grow these seeds. The contract sets an agreed price to be paid to the farmers for each kilogram of production. Farmers plant the seeds and once these seeds are matured, they start harvesting and continue until the end of the harvest season. Each week, the intermediary firm makes a payment to the farmers based on the total output from in that week, transport the quantity that is needed to satisfy the demand that week to a packaging facility, package them, and distribute to the retailers to meet their demand. In this agribusiness supply chain, the intermediary firm can realize a substantial benefit by managing the risks in the best way possible to match supply and demand.

Now consider one of possible farms for contracting located in Southern California. The farmer collected detailed information about the harvest time, maturation time, and the yield for different varieties of tomatoes (for example, see Chester 2011). Figure 5 shows the maturation time and harvest length for tomatoes in different years. Figure 6 gives the cumulative output from each plant. These figures show that the maturation time, harvest length, and the output exhibit high variability within a year and also from one year to another. Since an agribusiness works with a high number of contracted farms, they must be selected; the farm areas to be contracted must be determined; and the seeding times must be set in such a way that these variabilities are taken into account in the best way possible to maximize the expected profit. The cumulative output data can be used to estimate the farm availability probabilities and also the yield distribution parameters.

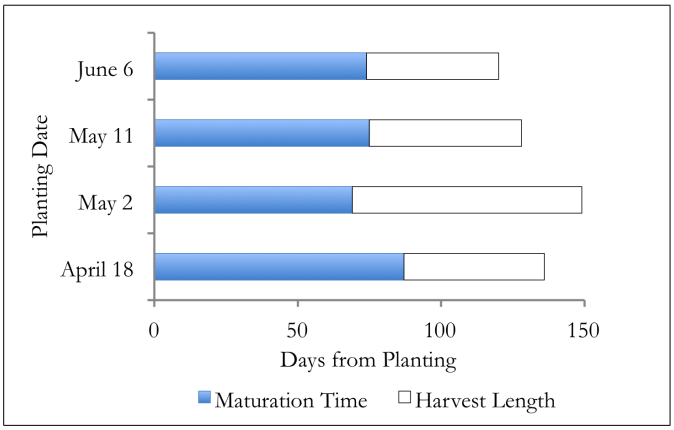


Figure 5 **Maturation Time and Harvest Length Observations at a Tomato Farm**

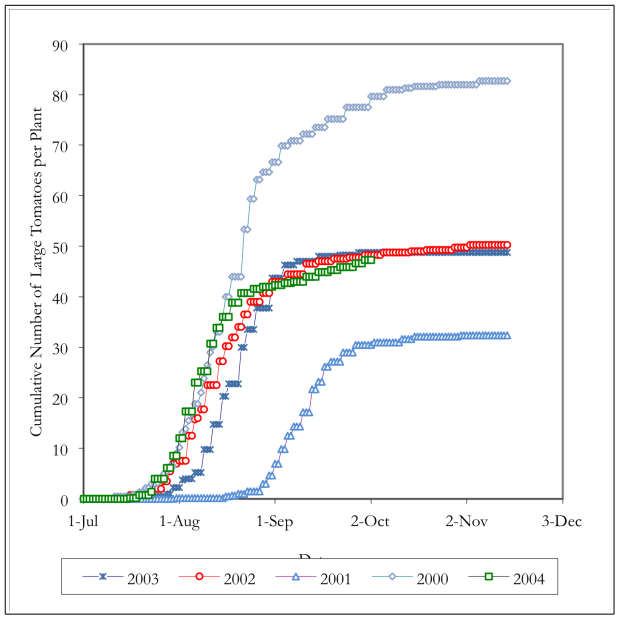


Figure 6 **Cumulative Output at a Tomato Farm in Different Years**

8.2. Performance of the Solution Methodology

We now consider a 10-period problem with up to 10 farms. In order to evaluate the performance of the solution methodology, we compare the solutions obtained by using the methodology presented in this paper with the approaches that do not consider variability. In order to test the performance in a wide range of system parameters, three demand patterns including a level demand, increasing, leveling, and then decreasing demand, and highly varying demand are considered for the expected values of the demand. For each pattern, low variability and high variability cases are considered. The expected maturation time is taken as 3 weeks and 4 different demand distributions are used.

Similarly, the expected harvest length is taken as 4 weeks and 4 different distributions are used. These combinations yield 980 different cases of the problem by using the parameters of scenarios given in Table 2.

Parameters	**Values**
N :	$1, ..., 10$
$\mathcal{T}$:	10
$r(t)$:	5
$c(t)$:	$2, 4$
$s(t)$:	0.5
$E[D(t)]$	(2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000)
	(2000, 2000, 3000, 3000, 2000, 2000, 2000, 2000, 2000, 2000)
	(2000, 4000, 1000, 3000, 2000, 3000, 1000, 4000, 2500, 1500)
$cv[D(t)]$	$0.1, 0.3$
$E[Y_i(t)]$:	6000
$cv[Y_i(t)]$:	0.1
M_{i,τ_i} :	Uniform(2,4), Uniform(1,5)
	$DD(1, 2, 3, 4, 5, ..., 10 : 0, 0.2, 0.6, 0.2, 0, ..., 0)$
	$DD(1, 2, 3, 4, 5, ..., 10 : 0, 0.2, 0.6, 0.2, 0, ..., 0)$
H_{i,τ_i} :	Uniform(3,5), Uniform(2,6)
	$DD(1, 2, 3, 4, 5, 6, ..., 10 : 0, 0, 0.2, 0.6, 0.2, 0, ..., 0)$
	$DD(1, 2, 3, 4, 5, 6, ..., 10 : 0, 0, 0.2, 0.6, 0.2, 0, ..., 0)$
	$DD(t_1, t_2, ..., t_{\mathcal{T}} : p_1, p_2, .., p_{\mathcal{T}})$ is a general discrete distribution.

Table 2 Parameters of Scenarios.

In order to evaluate the performance of the proposed method for the multi-farm multi-period problem for annual plants, we considered a number of benchmark solutions.

8.2.1. Comparison with the Exact Solution As a benchmark solution, first the optimal solution of the problem when $N \leq 4$ is determined for 288 scenarios generated by using the parameters given in Table 2. The farm areas that maximize the expected profit given in Equation (18) are determined for all possible combinations of seeding times to get the optimal solution for these cases.

Second, a mean-value solution is obtained by using a linear formulation for perennial plants that determines the farm areas when the seeding times are given and all the random variables in the original problem are replaced with their expected values. This formulation is given in Section 7.3. The mixed-integer linear formulation of the problem are solved by using CONOPT solver in GAMS to determine both the farm areas and the seeding times. This method is referred as AT_{exp}.

Finally, we find the solution by using a two-stage iterative method. For the numerical experiments, we start from an initial solution where all the farm areas are equal to each other and the seeding times are the ones that maximize the problem in (19)-(17) for equal farm sizes. The iterative solution procedure is continued until the difference between two successive solutions is smaller than a predetermined threshold. In the numerical experiments, the iterations are continued until the maximum deviation between the farm areas at the successive iterations for a given farm is less than 0.1%. It is observed that this procedure is quite efficient and converges to a solution in a few number of iterations. However, a formal proof for the convergence of the algorithm is not given in this paper. This method is referred as $AT_{norm-iter}$.

The farm areas and the seeding times obtained from each method is used to determine the expected total profit $E[\pi(\mathbf{a}, \tau)]$ by evaluating Equation (18). The average optimality gap that is the percentage difference between the optimal solution and the solution obtained by each method is shown in Table 3 when $N \leq 4$. Table 3 shows that modeling all the uncertainties explicitly is very beneficial. The model that considers all the uncertainties ($AT_{norm-iter}$) provides results that are 5.5% closer to the optimal solution compared to the model that only uses the expected values of all the random variables (AT_{exp}).

N	**Optimality Gap(%)**	
	$AT_{norm-iter}$	AT_{exp}
2	2.76	3.50
3	4.34	10.05
4	3.29	13.37
Average	3.46	8.97

Table 3 Comparison of different solution methods for the Multi-farm Multi-period problem for annual plants

8.2.2. Comparison with the Mean-value Solution Since finding the exact solution is not computationally feasible for large values of N, we compared the performance of the proposed iterative method to the solution obtained by using the expected values when $1 \leq N \leq 10$. 98 scenarios are generated for each N value by using the parameters given in Table 2. The percentage improvement of the proposed solution compared to the deterministic solution is depicted in Figure

7. The average improvement of these cases is around 15% and the proposed solution yields an expected profit that is more than 20% better compared the the mean-value approach when $N \geq 6$. These numerical experiments show that the proposed method performs substantially better

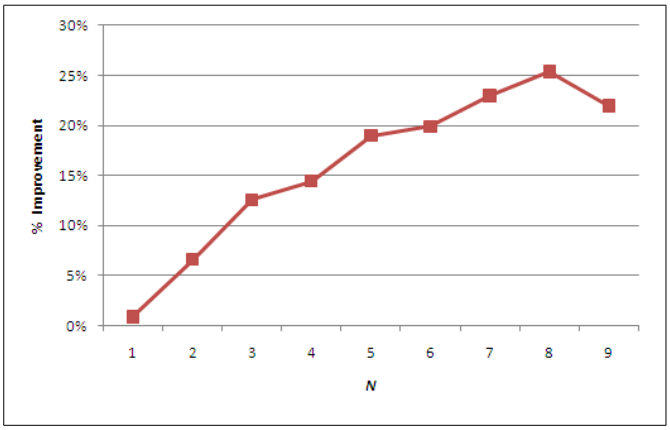


Figure 7 **Percentage improvement of the proposed solution over the mean-value solution for the Multi-farm Multi-period problem for annual plants**

compared to the planning approaches where the uncertainties are not taken into consideration directly. Moreover the comparative benefit of the stochastic planning approach gets better as the number of farms increases.

9. Conclusion

In this study, we consider an agricultural planning problem of a firm that contracts different farms located in different regions when both demand and supply are random. We first provide a detailed model of the supply from different farms that have maturation time, harvest time, and yield uncertainty. The single period problem is similar to the newsvendor problem with random yield. We provide a general solution to the newsvendor problem when the supply and the demand are normally distributed random variables. By using this result, we provide the analytical solution for the single-period single-farm case. This solution yields a nonlinear equation to be solved for the optimal area. The optimal seeding time is shown to be the one that gives the highest farm availability probability at the period where the demand is realized. For the single period case, we show that the seeding area decision can be separated from the seeding time decision. More

specifically, the maturation and harvest time uncertainty can be mitigated by determining the seeding time while the yield and demand uncertainty can be mitigated by determining the farm area.

The computational complexity to determine the objective function exactly increases exponentially with the number of farms. Therefore we present an approximation method that assumes normality of the total supply coming from multiple farms. For the single-period problem, our numerical experiments show that this approximation is very accurate and provides results that deviate from the optimal optimal solution by 2.3% on average for a single farm and improves very rapidly as as the number of farms increases. For example, the optimality gap decreases to 0.001% when the number of suppliers is 8.

In the second part, we analyze the multi-period, multi-supplier problem for annual plants. We propose an iterative approach to determine both the optimal farm areas and also the optimal seeding areas with the normal approximation for the total supply. We evaluate the performance of the solution methods by using a wide range of system parameters. The proposed iterative approach that determines both the farm areas and the seeding times yields a profit that is 16% higher on average than the profit obtained by using the mean-value approach that does not take uncertainty into account.

The realized increase in expected profit that is obtained by incorporating harvest, yield, and demand uncertainty directly in the planning method is quite substantial for the agricultural sector. Therefore we present this planning methodology as an effective tool for agricultural planning under yield, harvest, and demand uncertainty.

References

Ahumada, O., J.R. Villalobos. 2009. Application of planning models in the supply chain of agricultural products: A review. *European Journal of Operational Research* **196**(1) 1–20.

Ahumada, O., J.R. Villalobos. 2011a. Operational model for planning the harvest and distribution of perishable agricultural products. *International Journal of Production Economics* **133**(2) 677–687.

Ahumada, O., J.R. Villalobos. 2011b. A tactical model for planning the production and distribution of fresh produce. *Annals of Operations Research* **190** 339–358.

Allen, S.J., E.W. Schuster. 2004. Controlling the risk for an agricultural harvest, manufacturing. *Manufacturing and Service Operations Management* **6**(3) 225–236.

Anupindi, R., R. Akella. 1993. Diversification under supply uncertainty. *Management Science* **39**(8) 944–963.

Cain, M. 1994. The moment-generating function of the minimum of bivariate normal random variables. *The American Statistician* **48**(2) 124–125.

Çömden, E.N. 2009. Agricultural planning under demand, yield, and harvest uncertainty. Master's thesis, Koç University Industrial Engineering, Sariyer, Istanbul, Turkey.

Chester, T. 2011. Tomato harvest versus planting date. available from `http://http://tchester.org/analysis/tomatoes/index.html`, data received with personal communication.

Dada, M., N.C. Petruzzi, L.B. Schwarz. 2007. A newsvendor's procurement problem when suppliers are unreliable. *Manufacturing and Service Operations Management* **9**(1) 9–32.

Ferrer, J.C., A. MacCawley, S. Maturana, S. Toloza, J. Vera. 2008. An optimization approach for scheduling wine grape harvest operations. *International Journal of Production Economics* **112**(2) 985–999.

Gerchak, Y., R.G. Vickson, M. Parlar. 1988. A periodic review production models with variable yield and uncertain demand. *IIE Transactions* **20**(2) 144–150.

Itoh, T., H. Ishii, T. Nanseki. 2003. A model of crop planning under uncertainty in agricultural management. *International Journal of Production Economics* **81-82** 555–558.

Jones, P.C., G. Kegler, T.J. Lowe, R.D. Traub. 2003. Managing the seed-corn supply chain at Syngenta. *Interfaces* **33**(1) 80–90.

Kazaz, B. 2004. Production planning under yield and demand uncertainty with yield-dependent cost and price. *Manufacturing and Service Operations Management* **6**(3) 209–224.

Maatman, A., C. Schweigman, A. Ruijs, M.H. van der Vlerk. 2002. Modeling farmer's response to uncertain rain fall in burkina faso: A stochastic programming approach. *Operations Research* **50**(3) 399–414.

Merrill, J. M. 2007. Managing risk in premium fruit and vegetable supply chains. Master of Engineering in Logistics Thesis, Massachusetts Institute of Technology, Cambridge MA, USA.

Rekik, Y., E. Sahin, Y. Dallery. 2007. A comprehensive analysis of the newsvendor model with unreliable supply. *OR Spectrum* **29**(2) 207–233.

Romero, C. 2000. Risk programming for agricultural resource allocation: A multidimensional risk approach. *Annals of Operation Research* **94**(1-4) 57–68.

Schilizzi, S.G.M., R.S. Kingwell. 1999. Effects of climatic and price uncertainty on the value of legume crops in a mediterranean-type environment. *Agricultural Systems* **60**(1) 55–69.

Tan, B. 2011. *Managing Supply Chains on the Silk Road: Strategy, Performance, and Risk*, chap. Incorporating Harvest, Maturity, Yield, and Demand Risk in Planning for Agricultural Supply Chains for Premium Products. CRC Press, 239–252.

Widodo, K.H., H. Nagasawa, K. Morizawa, M. Ota. 2006. A periodical flowering-harvesting model for delivering agricultural fresh products. *European Journal of Operational Research* **170**(1) 24–43.

Yang, S., J. Yang, L. Abdel-Malek. 2007. Sourcing with random yields and stochastic demand: A newsvendor approach. *Computers and Operations Research* **34**(12) 3682–3690.

Yano, C.A., H.L. Lee. 1995. Lot sizing with random yields: a review. *Operations Research* **43**(2) 311–334.

Appendix

Proof of Theorem 1

The first derivative of $E[\pi(x)]$ given in Equation (7) is

$$\frac{\partial E[\pi(x)]}{\partial x} = (r-s)\left(\frac{\partial f(x)}{\partial x}\Phi\left(\frac{E[D]-f(x)}{h(x)}\right) - g(x)\frac{\partial g(x)}{\partial x}h(x)^{-1}\phi\left(\frac{E[D]-f(x)}{h(x)}\right)\right) - (c-s)\frac{\partial f(x)}{\partial x} \tag{23}$$

where $h(x) = \sqrt{g^2(x) + Var[D]}$.

The second derivative of the profit function can be written as

$$\frac{\partial^2 E[\pi(x)]}{\partial x^2} = \frac{\partial^2 f(x)}{\partial x^2}\varphi_1 - h(x)^{-3}\varphi_2\phi\left(\frac{E[D]-f(x)}{h(x)}\right)$$

where $\varphi_1 = (r-s)\,\Phi\left(\frac{E[D]-f(x)}{h(x)}\right) - (c-s)$ and

$$\varphi_2 = \left(\frac{\partial f(x)}{\partial x}h(x) + g(x)\frac{\partial g(x)}{\partial x}\left(\frac{E[D]-f(x)}{h(x)}\right)\right)^2 + Var[D]\left(g(x)\frac{\partial^2 g(x)}{\partial x^2} + \left(\frac{\partial g(x)}{\partial x}\right)^2\right) + g(x)^3\frac{\partial^2 g(x)}{\partial x^2}.$$

The terms $\Phi\left(\frac{E[D]-f(x)}{h(x)}\right)$ and $\phi\left(\frac{E[D]-f(x)}{h(x)}\right)$ are always non-negative. Also $h(x) = \sqrt{g^2(x)+Var[D]}$ and $g(x)$ that is the standard deviation of the total supply are non-negative. Since $r > c$ but $0 \leq \Phi(\cdot) \leq 1$, φ_1 can be positive or negative.

If $\frac{\partial^2 g(x)}{\partial x^2} \geq 0$ then φ_2 is guaranteed to be non-negative. Then if $\frac{\partial^2 g(x)}{\partial x^2} \geq 0$, and $\frac{\partial^2 f(x)}{\partial x^2} = 0$, $\frac{\partial^2 E[\pi(x)]}{\partial x^2}$ is guaranteed to be non-positive. Equivalently, if $g(x)$ is convex and $f(x)$ is linear, then $E[\pi(x)]$ is concave and x value that satisfies the first order optimality condition is the optimal x that maximizes $E[\pi(x)]$. By using Equation (23), rearranging $\frac{\partial E[\pi(x)]}{\partial x} = 0$ yields the equality given in Equation (8) $\square$

Proof of Theorem 2

The structure of the expected profit for period t^* given in Equation(10) is the same as the expected profit function for the newsvendor problem with random supply given in Equation (7). Since $E\left[Q'(t^*)\right] = f'(a_1) = E\left[Y_1(t)\right] a_1$ is linear and $\sigma\left[Q'(t^*)\right] = g'(a) = \sigma\left[Y_1(t)\right] a_1$ is also linear, the objective function is concave according to Theorem 1. Furthermore the optimal a_1 must satisfy Equation (8):

$$f'(a_1) = E\left[D(t^*)\right] - \theta'(t^*)\Phi^{-1}\left[\frac{(c(t^*)-s(t^*))}{(r(t^*)-s(t^*))} + g'(a_1)\frac{\partial g'(a_1)}{\partial a_1}\theta'(t^*)^{-1}\phi\left(\frac{E\left[D(t^*)\right]-f'(a_1)}{\theta'(t^*)}\right)\left(\frac{\partial f'(a_1)}{\partial a_1}\right)^{-1}\right] \tag{24}$$

After replacing the $f'(a_1)$, $g'(a_1)$, $\frac{\partial g'(a_1)}{\partial a_1}$ and $\frac{\partial' f(a_1)}{\partial a_1}$ in Equation (24), the optimality condition for a_1 becomes the one given in Equation (11). $\square$

Proof of Theorem 3

The derivative of the expected profit given in Equation (10) with respect to $p_{1,\tau_1}(t^*)$ is

$$\begin{aligned}\frac{\partial E[\pi\left(a_1,\tau_1\right)]}{\partial p_{1,\tau_1}(t^*)} = {} & (r(t^*)-s(t^*))\Bigg(E\left[Q'(t^*)\right]\Phi\left(\frac{E\left[D(t^*)\right]-E\left[Q'(t^*)\right]}{\theta'(t^*)}\right) \\ & +E\left[D(t^*)\right]\Phi\left(\frac{E\left[Q'(t^*)\right]-E\left[D(t^*)\right]}{\theta'(t^*)}\right) - \theta'(t^*)\phi\left(\frac{E\left[D(t^*)\right]-E\left[Q'(t^*)\right]}{\theta'(t^*)}\right)\Bigg) \\ & -(c(t^*)-s(t^*))E[Q(t^*)']\end{aligned}$$

which can be rewritten as

$$\frac{\partial E[\pi\left(a_1,\tau_1\right)]}{\partial p_{1,\tau_1}(t^*)} = \frac{E[\pi\left(a_1,\tau_1\right)]}{p_{1,\tau_1}(t^*)}.$$

It is known that in the optimal solution $E[\pi(a_1, \tau_1)] \geq 0$ and $p_{1,\tau_1}(t^*) \geq 0$. Then $\frac{\partial E[\pi(a_1,\tau_1)]}{\partial p_{1,\tau_1}(t^*)} \geq 0$.

Let us consider a single farm that has the farm area of $a_1 > 0$ that maximizes the expected profit given in Equation (10) when the seeding time is τ_1. Let $\tau_1' \neq \tau_1$ be another seeding time, such that the farm availability probability at period t^* given that the seeding time is $\tau_1, p_{1,\tau_1}(t^*), is$ greater then the same probability when the seeding time is $\tau_1', p_{1,\tau_1'}(t^*)$. Since $\frac{\partial E[\pi(a_1,\tau_1)]}{\partial p_{1,\tau_1}(t^*)} \geq 0$, $E[\pi(a_1, \tau_1)] \geq E[\pi(a_1, \tau_1')]$ if and only if $p_{1,\tau_1}(t^*) \geq p_{1,\tau_1'}(t^*)$. If τ_1 is the optimal seeding time, $E[\pi(a_1, \tau_1)] \geq E[\pi(a_1, \tau_1')]$ for $\forall \tau_1' \neq \tau_1$. As a result, the optimal seeding time is the time that yields the highest farm availability probability at the desired period t^*. □

Proof of Theorem 4

Equation (13) can be rewritten as

$$E[\pi(\mathbf{a}, \tau)] = \sum_{\alpha_1=0}^{1} ... \sum_{\alpha_N=0}^{1} \left(\prod_{i=1}^{N} p_{i,\tau_i}^{\alpha_i}(t^*) \left(1 - p_{i,\tau_i}(t^*)\right)^{1-\alpha_i} \right) E[A(\mathbf{a}, \tau)]$$

where

$$\begin{aligned} A(\mathbf{a}, \tau) = & \left(r(t^*) - c(t^*)\right) \left(\sum_{i=1}^{N} \left(E[Y_1(t^*)] + Z\sigma[Y_{1,t^*}]\right) \right) a_i \alpha_i \\ & - \left(r(t^*) - s(t^*)\right) \left(\sum_{i=1}^{N} \left(E[Y_1(t^*)] + Z\sigma[Y_1(t^*)]\right) a_i \alpha_i - D(t^*) \right)^+ \end{aligned}$$

and Z is a standard normal random variable.

The function $A(\mathbf{a}, \tau)$ can be represented as the limit of a sequence $\{A_m\}$. In $A_m(\mathbf{a}, \tau)$, the random variables Z and $D(t^*)$ take values with corresponding probabilities in a finite range. Since $(\cdot)^+$ is a convex function, for each pair of the values of Z and $D(t^*)$, $A_m(\mathbf{a}, \tau)$ is concave in $\mathbf{a}$. Then $E[A_m(\mathbf{a}, \tau)]$ is also concave. Since the sequence $\{A_m\}$ converges pointwise to $A(\mathbf{a}, \tau)$, $E[A(\mathbf{a}, \tau)]$ is also concave. Finally $E[\pi(\mathbf{a}, \tau)]$ is a linear combination of concave functions and therefore it is also concave in $\mathbf{a}$. □

Proof of Theorem 6

Let us consider a farm whose farm area a_i is set to maximize the the expected profit function given in Equation (14) for the given seeding time τ_i. The derivative of the expected profit function given in Equation (14) with respect to p_{i,τ_i} can be written as

$$\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial p_{i,\tau_i}} = a_i\Bigg(E[Y_i(t)]((r(t)-c(t))\Phi(\tfrac{E[D(t)]-E[Q(t)]}{\theta(t)}) - (c(t)-s(t))) - \\ -(r(t)-c(t))\tfrac{a_i(E[Y_i(t)^2]-2p_{i,\tau_i}(t)E[Y_i(t)]^2)}{2\theta(t)}\phi(\tfrac{E[D(t)]-E[Q(t)]}{\theta(t)})\Bigg)$$

Since a_i is set to maximize the expected profit, $\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial a_i} = 0$. By reorganizing Equation (15),$\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial a_i}$ can be written as follows;

$$\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial a_i} = p_{i,\tau_i}\Bigg(E[Y_i(t)]((r(t)-c(t))\Phi(\tfrac{E[D(t)]-E[Q(t)]}{\theta(t)})\theta(t)) - (c(t)-s(t))) - \\ (r(t)-c(t))\tfrac{a_i(E[Y_i(t)^2]-p_{i,\tau_i}(t)E[Y_i(t)]^2)}{\theta(t)}\phi(\tfrac{E[D(t)]-E[Q(t)]}{\theta(t)})\Bigg) = 0.$$

In both equations, $r(t)-c(t)>0$, $a_i>0$, $\theta(t)>0$ and $\phi\left(\frac{E[D(t)]-E[Q(t)]}{\theta(t)}\right)>0$ by definition. Since $Var[X]=E[X^2]-E^2[X]\geq 0$ and $0\leq p_{i,\tau_i}(t)\leq 1$, $E[Y_i(t)^2]-2p_{i,\tau_i}(t)E[Y_i(t)]^2\geq 0$ and $E[Y_i(t)^2]-p_{i,\tau_i}(t)E[Y_i(t)]^2\geq 0$. Since the first terms are the same in the two equations,to show $\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial p_{i,\tau_i}}\geq 0$ while $\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial a_i}=0$, it is sufficient to prove

$$(r(t)-c(t))\frac{a_i(E[Y_i(t)^2]-p_{i,\tau_i}(t)E[Y_i(t)]^2)}{\theta(t)}\phi\left(\frac{E[D(t)]-E[Q(t)]}{\theta(t)}\right)\geq \\ (r(t)-c(t))\frac{a_i(E[Y_i(t)^2]-2p_{i,\tau_i}(t)E[Y_i(t)]^2)}{2\theta(t)}\phi\left(\frac{E[D(t)]-E[Q(t)]}{\theta(t)}\right)$$

The difference of these two terms can be written as $\frac{a_i(r(t)-c(t))}{2\theta(t)}\phi(\frac{E[D(t)]-E[Q(t)]}{\theta(t)})E[Y_i(t)^2]$. Since all multiplier factors are greater or equal to zero, the term is non-negative. This proves that when $\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial a_i}=0$, $\frac{\partial E[\tilde{\pi}(\mathbf{a},\tau)]}{\partial p_{i,\tau_i}}\geq 0$.

Let us consider two seeding times τ_i^k and $\tau_i^{k'}$, such that the probability of harvesting in period t given that seeding time is $\tau_i^k, p_{i,\tau_i^k}(t)$, greater then the same probability when the seeding time is $\tau_i^{k'}, p_{i,\tau_i^{k'}}(t)$. Let $\mathbf{a}_i^{k*}$ denote the optimal $\mathbf{a}$ for the seeding time τ_i^k. The above result shows that for all τ_i^k,$\tau_i^{k'}\in\mathcal{T}$, $\tilde{\pi}\left(\mathbf{a}_i^{k*},\tau_i^k\right)$ is greater or equal to $\tilde{\pi}\left(\mathbf{a}_i^*,\tau_i^{k'}\right)$ when $p_{i,\tau_i^1}(t)\geq p_{i,\tau_i^2}(t)$. As a result the seeding time that maximizes the expected profit under the normality assumption of the total supply is the one that yields the highest farm availability probability $\square$.