EVALUATION OF SIMULATION-BASED OPTIMIZATION IN GRAFTING LABOR ALLOCATION

S. Masoud, Y. J. Son, C. Kubota, R. Tronstad

ABSTRACT. *Vegetable grafting is a labor-intensive operation with many management decisions. Labor management and resource planning are critical allocations in grafting nurseries, yet optimization is challenging due to the dynamic nature of workers' performance in vegetable seedling propagation. To this end, we developed a simulation-based optimization framework for labor management to optimize labor allocation. This approach was evaluated by comparing its result with those suggested by selected domain experts (a plant scientist and a nursery manager). Furthermore, the simulation models were validated with a dataset from a developing tomato grafting company. Simulation-based optimization is demonstrated as an effective approach to find the optimal/near optimal labor allocation for horticultural nurseries, where discrete event simulation is used to represent the dynamics of the grafting work environment and meta-heuristics are used to devise optimal/ near optimal resource allocation strategies. Results reveal that a potential annual savings between \$2,510 (0.6%) and \$97,388 (20%) can be achieved for a grafting facility of 6 million plants per year.*

Keywords. *Simulation-based optimization, Grafting, Labor allocation, Discrete event simulation.*

Vegetable grafting can improve crop yields by acquiring resistance against diseases and pests (Kubota et al., 2008; Guan et al., 2012). In grafting, a plant that has greater yield and/or more desirable fruit properties is used for the upper or harvestable plant part (i.e., scion), while a plant that is more resistant to pests is used for the lower part that roots into the ground (i.e. rootstock) (Lee et al., 2010). As a result, grafted plants are expected to be not only more productive compared to non-grafted scion plants, but also more tolerant against both biotic and abiotic stresses (Chiu et al., 2008; Schwarz and Colla, 2010; Schwarz et al., 2010; Louws et al., 2010; Guan et al., 2012).

Despite the biological advantages of grafting in vegetable production, the process is labor-intensive and costly. The grafting process consists of four main tasks: 1) cutting the scion seedling at a specific angle, 2) cutting the rootstock seedling at the same angle, 3) placing a clip of the right size over the cut end of rootstock seedling, and 4) firmly joining the scion and rootstock together. Each of these tasks requires a high level of workers' expertise and sometimes specific equipment to enhance both speed and success rate (Meng

et al., 2014). Although plant health and compatibility of scion and rootstock are key factors for successful grafting, the role of human capital is also significant. The size of a grafting operation varies per production scale, requiring from a few workers to more than 100 workers. When workers cut the scion and rootstock seedlings to join them, precise alignment of two cut ends of seedlings is required as the alignment of cambial layers supports the transport of water within the new grafted plant (Stoddard and McCully, 1979).

Variations in human skills and speed may cause problems with precise cutting or alignment of end cuts, which will negatively affect the success rate of grafting. These problems occur not only among different workers, but also within a given worker. While automation may be considered as a solution to resolve this problem of variation, the large capital investment of automation cannot be easily justified (Meng et al., 2012). In addition, specialized automated machinery cannot be moved to other jobs and they cannot be idle during the off-season (Meng, et al., 2014). Consequently, although automation has greatly advanced over the past 20 years, it has not been cost-effective for most small facilities due to its high capital investment and requirements for expensive materials (Kubota et al., 2017). Labor management plans are still critical for grafting facilities to improve their productivity since labor is the single highest cost for manual grafting (Calvin and Martin, 2011).

When a nursery manager designs logistics of the grafting operation, two major types of workflows are of concern. In the first workflow, each of four tasks discussed earlier (such as cutting scion or joining scion and rootstock) are designated to one or more specific worker, based on his/her skill level and experience. This workflow is known as the in-line

Submitted for review in June 2017 as manuscript number ITSC 12487; approved for publication by the Information, Technology, Sensors, & Control Systems Community of ASABE in January 2018.

The authors are **Sara Masoud**, Graduate Student and Young-Jun Son, professor, Department of Systems and Industrial Engineering, The University of Arizona, Tucson, Arizona; **Chieri Kubota**, Professor, School of Plant Sciences, The University of Arizona, Tucson, Arizona; and **Russell Tronstad**, Distinguished Outreach Professor and Specialist, Department of Agricultural and Resource Economics, The University of Arizona, Tucson, Arizona. **Corresponding author:** Young-Jun Son, 1127 E. James E. Rogers Way, Room 111, PO Box 210020, Tucson, AZ 85721-0020; phone: 520-626-9530; e-mail: son@sie.arizona.edu.

operation in manufacturing. A major advantage of this workflow is that each worker can focus on a single task that he/she is trained for. In the second workflow, each worker handles all four tasks, which is similar to cellular manufacturing. In this setting, each individual is responsible for each grafted tray, and thereby individual performance is easier to evaluate. A nursery manager typically selects one of these workflows based on his/her experiences.

In general, labor management for a manual grafting nursery is complex due to three major factors. First, workflow of the facility depends on performance of individual workers with various levels of experience. Second, absenteeism of workers due to personal reasons can increase the complexity of labor allocation (Andersson and Olsson, 1998). Third, seasonal demand for transplants makes the resource (i.e., both workers and other supplies) allocation of production lines complicated. In addition, the perishable nature of grafted seedlings makes labor management even more challenging and good management also helps buffer against needing additional investments like cold storage (Andersson and Olsson, 1998; Kubota et al., 2017). Uncertainty caused by these three factors affects the productivity of grafting lines, especially manual operations. To this end, our study focuses on optimizing labor resource allocation by considering the impact of the variability in workers' performance. At the same time, facilities are welcome to employ our proposed framework to find the optimal/ near optimal resource allocation for different demand patterns as well.

Simulation based optimization has been used for studying the impacts of uncertainty in resource management for manufacturing systems (Rogers, 2002; Creighton and Nahavandi, 2003; Kleijnen, 2015). Optimization has been utilized in dynamic and uncertain agricultural systems, where various decisions were considered such as evaluating economics, design, and plant growth (Meng et al., 2014; Rasouli et al., 2014; McCorkle et al., 2016; Kubota et al., 2017; Parthanadee and Buddhakulsomsiri, 2010). For example, Ferrer et al. (2008) proposed a mixed integer-linear programming model to find optimal harvest scheduling for wine grapes considering operating costs and quality. Allen and Schuster (2004) developed a robust mathematical model in order to optimize the risk of harvesting Concord grapes. Kaza (2004) presented a stochastic optimization model for production planning of olive oil considering both random yield and demand. Soto-Silva et al. (2016) provides a thorough survey of optimization models applied to the fresh fruit supply chain. Although these models have been utilized to consider the uncertainty rooted in demand, yield, or raw material in agricultural crops, to the best of our knowledge, no research has been conducted to study the impact of uncertainty caused by workers' performance, especially in a labor-intensive horticultural operation like vegetable grafting. The application of optimization models utilized in different crops as noted has been less successful in this area due to its lack on capturing the uncertainty of the system and inherent complexity (Bohle et al., 2010). For example, modeling the growth of a plant as a closed-form equation in an optimization model is not easy. For this study, it is not easy to develop closed-form equations to mimic the random growth level of

two different plant types (i.e., scions and rootstocks), the sequential process of grafting, and the relation between these processes and workers' skill levels. In such situations where the complexity of the system cannot be fully incorporated by optimization involving closed-form equations, simulation based optimization is a viable solution. An additional benefit of utilizing a simulation model for grafting is its ability to provide output analysis based on various levels of uncertainty as proposed by Better et al. (2015).

The main goal of this study is to improve the efficiency of a grafting nursery operation through optimizing labor allocations by considering the variety of workers performance. In this work, we utilize discrete event simulation to model detailed grafting operations of a selected commercial nursery, and to compare various workflows. In general, discrete event simulation defines the operations of a system as a sequence of events in time, where each task (e.g., cutting a single scion, or grafting a plant) is occurring at a specific time stamp related to the event. Each event (e.g., changing state from cut rootstock to clipped rootstock) takes place at a particular instant in time, and no change can occur between consecutive events; thus, the simulation can jump in time between events. As a result, the simulation time proceeds faster in comparison to real wall-clock time. In addition, these simulation models are employed in our simulation based optimization framework to help optimize the labor allocation for improved performance and positive economic impacts. For each task, workers with diverse levels of expertise (i.e., high, moderate, and low experienced workers) are considered. The simulation models constructed in this work incorporate all aspects of operations, such as processing time, success rate, rework processing time, and material handling under different configurations. Afterward, we evaluate solutions suggested by field experts and compare them with the optimized solution to discuss further the impact of different labor configurations in nursery labor management for each workflow.

METHODS

The dynamics of a grafting operation are complex due to varying levels of skills among workers and within each worker, therefore these dynamics cannot be adequately represented by a closed-form analytical model. To address this challenge, we propose simulation based optimization to find the optimal/ near optimal labor allocation for grafting facilities. To evaluate each workflow, we collect suggested solutions by field experts and compare them with the solutions suggested by the simulation based optimization (i.e., case study). Our approach includes four major steps: 1) data collection and analysis of the statistical distributions, 2) case study by domain experts, 3) development of simulation based optimization, and 4) comparison of solutions. We also validated the developed models by comparing the model outputs with the actual data obtained in the considered company.

DATA COLLECTION AND ANALYSIS OF STATISTICAL DISTRIBUTION

Figure 1 displays the performed processes in order to identify the statistical distribution of processing time in this

Figure 1. Data gathering and analysis is conducted to sample the workers' performance and identify the best processing time distribution via Kolmogorov-Smirnov, Anderson-Darling, and Chi-square tests.

work. In particular, processing times of the four grafting operation tasks (i.e., scion cutting, rootstock cutting, rootstock clipping, and joining) were collected at a commercial tomato grafting facility with production capacity of 7200 plants per day. This data set of 70 data points (i.e., 3 groups of at least 23 samples for each level of expertise) guarantees a power of 95% at $\alpha=0.05$ statistical significance. Utilizing this dataset, parameters were fit for processing time distributions of each grafting operation task. Here, three different goodness of fit tests (i.e., Kolmogorov-Smirnov, Anderson-Darling, and Chi-squared) were conducted to select the most suitable distribution for each task.

CASE STUDY BY DOMAIN EXPERTS

After identifying the processing time distributions, we asked our domain experts to utilize their knowledge and expertise to suggest the best labor configuration to minimize the labor costs for each workflow based on their experiences, when a nearly triple production capacity (i.e., 24,000 plants or 100 trays per day) was intended by hiring a new work force and allocating tasks for individuals in a selected workflow. They filled out a survey form, which contained the derived processing time distributions, alongside the list of assumptions based on which the experiments had been designed. Table 1 provides the mean and standard deviation (Std) of workers performance data regarding four main tasks by their experience.

In addition to workers performance data, the following conditions were provided to the domain experts and used in our simulation based optimization:

- Facility's objective is to produce 100 trays (i.e., 24,000 grafted plant) per day with minimum daily labor cost;
- Each worker's level of experience and performance (one of three levels defined in table 1) is known at the time of hiring;
- For in-line workflow, each potential worker is assigned to a specific task;
- In cellular workflow, each worker will handle all tasks related to grafting a plant; and
- Workers work 8 hours per day.

SIMULATION-BASED OPTIMIZATION

Simulation -based optimization is the process of integrating optimization methods and simulation models, where the simulation model evaluates the impacts of the parameters optimized by the optimization model, and updates the optimization model based on its given evaluation (Gosavi, 2003). Figure 2 provides the overall flow of the simulation-based optimization framework employed in our study.

These processes start with a labor allocation, which is an initial solution suggested by the optimization model. The suggested solution will be implemented in the simulation model. The mentioned process will continue until the optimal/near optimal solution is achieved. The simulation and optimization models which define the simulation based optimization framework are introduced in the following subsections.

Simulation Model

As mentioned earlier, we use discrete event simulation to model the processes of the under-study grafting company. The verification and validation of the simulation model are major tasks in any study involving simulation to ensure the accuracy and credibility of the simulation model. Verification is the process of ensuring that a model behaves as intended (Kelton et al., 2002; Fishman, 2013). To verify our simulation model in this study, we firstly allowed only a single grafting plant to enter the system and followed that entity until it left the system to ensure that the model logic and data were correct. Then we ran a few plants through the system using deterministic values for arrival and service times. We then calculated the predicted behavior by hand and verified that generated models matched that behavior exactly. Validation is the process of ensuring that a model reproduces the behavior of the real system. The validity of the simulation model depends on the accuracy level of input data and the modelling process. To ensure validity of input data, we used the data collected in the company, identified statistic distributions to represent the data, and used them as inputs for our simulation models in our study. For validity of simulation model, the current labor cost of the considered company having two workers of high skill level, two workers of moderate skill level, and one worker of low skill level was compared

Figure 2. Simulation based optimization framework where simulation evaluates various labor allocations, and obtains the best configuration.

Table 1. Workers skill based processing time for each task.

Level of Experience	Rootstock Cutting (min/tray)		Rootstock Clipping (min/tray)		Scion Cutting (min/tray)		Joining (min/tray)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Low	13.75	1.48	15.50	2.29	25.67	0.47	64.00	6.00
Moderate	10.15	2.74	13.69	2.97	17.64	5.66	50.67	5.25
High	8.56	1.26	12.89	2.88	16.44	4.14	43.39	6.36

with the simulation model output to ensure the company data is within the 95% confidence interval of the simulation output.

In addition to the assumptions described in the case study section by domain experts, this section provides additional information considered in the constructed simulation models. The simulation model for each workflow (i.e., in-line and cellular) consists of a set of breaks including a 30-min lunch break and two 15-min arbitrary breaks. Material flow between departments takes place via trays containing 240 plants, but single plants are modelled in each departmental process. The percentage of work in progress that needs rework depends on both the workflow and the process. With in-line workflow, after each process an inspection occurs to check the quality of the work, while in the cellular workflow the inspection occurs after the joining process only. If the quality of work does not pass inspection, the plant will go through a rework process that is defined for each task. In addition, for in-line workflow, the cutters and

clippers start the shift 45 min earlier, and consequently leave 45 min earlier. This step is defined as a solution to balance out the bottleneck of the joining process.

The simulation model, representing the in-line workflow, is designed in two levels (see fig. 3 for a snapshot of the simulation schema, developed in Arena software). The top-level controls the overall logic of the model such as the relationship between different departments (i.e., scion cutting, rootstock cutting, rootstock clipping, and joining) and labor availability check. The bottom levels represent individual departments, where each one is engaged in one of the mentioned particular tasks (i.e., scion cutting, rootstock cutting, rootstock clipping, and joining). In each department, workers with three different skill levels are modelled. They are assigned to work based on the suggested solution by each source of information (i.e., simulation based optimization, a plant scientist, and a nursery manager).

In this model of in-line workflow, each worker is assigned to a predefined single task. Each worker focuses on

Figure 3. A snapshot of simulation run for in-line, where each worker oversees a specific task.

finishing a whole tray. When a tray is completed, the worker will pass it to the next department for the next task. At the top level, trays are defined as active entities and when labor is available in each department and there is a need for production (i.e., the daily production target is not achieved yet), the trays will be sent to their respective departments.

Figure 3 displays one of the bottom level models which is the scion cutting department. This department is in charge of cutting scions. When a tray enters this department, it will be assigned to a worker with a specific skill level, who will start cutting the scion seedlings in the tray and check the quality of the cuts to decide if the seedling needs to be reworked. When the whole tray is completed, the same worker will pass the tray to the next department. Although figure 3 displays the process of scion cutting by a highly experienced worker only, those involving different skill levels follow similar logics and the only difference is in their processing time.

Figure 4 displays a snapshot of a simulation schema constructed in Arena software (Rockwell Automation Technologies, Pittsburgh, Pa.) to represent the cellular workflow, where each individual worker is in charge of all the tasks regarding the production of a grafted plant. In this model, like in-line workflow, workers with three various levels of expertise are available. Based on the configurations provided by the information sources, a predefined set of workers are available in the system as shown in table 3. Although the logic of modelling workers with different skill levels is the same, they have different processing times for the same task based on their skill level. Each worker will pick one tray of

both scion and rootstock seedlings and perform the mentioned four tasks plus inspection of each plant until the whole tray is done.

The main difference compared with the first model is that all workers can handle every task needed for grafting a plant. Here, each worker checks the quality of their work after finishing the joining process for each plant and if necessary, the worker spends additional time to rework the plant to improve the quality of the graft. In this model, workers are classified based on their level of expertise.

Optimization Model

Equations 1 to 7 represent the optimization model we implement. The main goal of this model is to minimize the total daily cost while guaranteeing the feasibility of the targeted production goal (i.e., M plants per day.).

$$\min w = \sum_i \sum_j c_i x_{ij} \quad (1)$$

Subject to:

$$grafted - cutScion \leq 0 \quad (2)$$

$$grafted - clippedRootStock \leq 0 \quad (3)$$

$$clippedRootStock - cutRootStock \leq 0 \quad (4)$$

$$grafted \geq M \quad (5)$$

Figure 4. A snapshot of simulation run for cellular, where a worker handles all tasks relating to grafting.

$$\sum_{i=1}^3 \sum_{j=1}^4 x_{ij} > 0 \quad (6)$$

$$x_{ij} \geq 0 \quad i \in \{1, 2, 3\}, j \in \{1, 2, 3, 4\} \quad (7)$$

where equation 1 is the objective function, which minimizes total daily labor costs. x_{ij} is the decision variable, denoting the number of workers with i^{th} level of expertise (i.e., high, moderate, and low) doing the j^{th} task (i.e., scion cutting, rootstock cutting, rootstock clipping, and grafting). c_i is the hourly wage of the workers, which is \$12.50, \$10.50, and \$8.50 for workers with high, moderate, and low levels of expertise, respectively. While minimizing daily labor costs, the optimization model ensures that daily demand is met via equations 2-4. *Grafted*, *cutScion*, *cippedRootStock*, and *cutRootStock* are the response variables. These response variables do not appear in the objective function and may seem to be irrelevant to decision variables (i.e., x_{ij}), but these variables (i.e., response variables) bring together the simulation model and the optimization model (Laguna, 1997; Wan and Li, 2008). In other words, decision variables are the variables that the modeler has control over and wants to optimize, but the response variables are the ones that measure the status of the simulation model in response to different values of the decision variables. Here, during each iteration, the decision variables (i.e., x_{ij}) are set by OptQuest software (utilizing meta-heuristics) based on the given setting and the simulation model runs in Arena software. While running, the response variables (i.e., *grafted*, *cutScion*, *cippedRootStock*, and *cutRootStock*) evaluate the suggested setting (i.e., x_{ij}). If the given labor allocation is not feasible to produce M trays of grafted plants, in the next iteration OptQuest will try to find a feasible labor allocation. Otherwise OptQuest will assign a new set of decision variables to improve the solution (i.e., reducing the daily labor cost) until the optimal/ near optimal solution is achieved.

Equations 2 and 3 assure that at least M cut scions and clipped rootstock trays are available to produce M grafted plants. Equation 4 assures the same for the clipped and cut rootstocks. Equation 5 represents the production goal (e.g., 100 trays of grafted plants) of the grafting facility. Although the constraint is relaxed to be greater or equal to M , the minimization objective will push the number of trays consisting of the grafted plants to M . Equation 6 ensures that at least one worker is available in the system. Finally, equation 7 ensures that the number of workers is a positive integer. These models are implemented in Arena and OptQuest software to achieve simulation based optimization.

For the cellular workflow (i.e., the one that each worker handles all the tasks for grafting a plant) only equations 1 and 5 are relevant. One additional equation (i.e., $x_{ij} \in \{0, 1\}$) is defined for each worker j with the expertise level of i . Here, x_{ij} is a binary variable to denote whether the j^{th} (where $0 < j < 20$) worker with the i^{th} level of expertise will be assigned (i.e., $x_{ij}=1$) in the optimization model.

COMPARISON OF SOLUTIONS

The last step is to compare the solutions that were suggested by the three information sources. In particular, considering an average of 250 working days (i.e., 365 days – 103 weekend days – 12 public holidays) during the year, we utilized the simulation models that were used for simulation based optimization in order to have a standard platform for comparing the impact of the suggested solutions. For a specific annual production of 6 million plants (i.e., 24,000 plants/day $\times$ 250 work days/year), each information source suggested a resource allocation for both in-line and cellular configurations. Those suggested configurations are explained in the following section. After getting the advised configurations from field experts, the suggested configurations were evaluated via simulation models. Then, the results of simulation models for different workflows were compared.

RESULTS AND DISCUSSION

As mentioned, the main goal of this study is to propose a simulation-based optimization framework for labor management in grafting nurseries. We evaluated performance of the developed framework by comparing the result (i.e. labor allocation) obtained by the framework with those suggested by field experts.

The first step, as discussed, is to obtain best-fit statistical distribution of processing time for each predefined task. Based on the data collected at a manual tomato grafting facility, distributions of processing times were obtained for major operations at the facility: 1) cutting a tray of scions (17.64 ± 5.66 min/tray), 2) cutting a tray of rootstock (i.e., 10.15 minutes/tray and ± 2.74 min/tray), 3) clipping a tray of rootstock (13.69 ± 2.97 min/tray), and 4) grafting the trays of cut scion and clipped rootstock (i.e., 46.63 min/tray ± 9.00 min/tray). Although each tray consists of 240 cells, scion trays have been utilized as double seeded. As a result, each scion tray is supporting 480 scion seedlings. Three different goodness of fit tests: 1) Kolmogorov-Smirnov, 2) Anderson-Darling, and 3) Chi-squared were utilized and the results of these tests are shown in table 2 for mid-level experienced workers.

The hypothesis for tests in table 2 involves checking if the distributions of the processing times for each task (i.e., scion cutting, rootstock cutting, rootstock clipping, and joining) are normally distributed with the given parameters (see the third row in table 1). As shown in table 2, all the calculated statistics are less than the mentioned critical values, therefore, the null hypothesis cannot be rejected. Thus, it's reasonable to utilize normal distributions for the processing times of the tasks in our analyses. Similarly, the processing time distributions for high-level experienced and low-level experienced workers were similarly identified. These distributions are used in the constructed simulation model.

Given the estimated distributions, we asked our field experts to allocate workers based on the mentioned assumptions. And by running the optimization model via OptQuest software over the simulation models, we get the simulation based optimization solutions as well. The solutions from all

Table 2. Kolmogorov-Smirnov, Anderson-Darling, and Chi-Squared goodness of Fit tests to evaluate the distribution of the gathered data.

Null hypothesis: Each task follows a normal distribution with the given mean and standard deviation.

	Scion Cutting (Min/Tray)				Rootstock Cutting (Min/Tray)				Rootstock Clipping (Min/Tray)				Grafting (Min/Tray)			
	Mean: 17.64	Std: 5.66			Mean: 10.15	Std: 2.74			Mean: 13.69	Std: 2.97			Mean: 46.63	Std: 9		
Kolmogorov-Smirnov																
Statistic	0.1707				0.1955				0.1694				0.1496			
P-Value	0.7488				0.6346				0.792				0.7345			
A	0.1	0.1	0.02	0.01	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01
Critical Value	0.34	0.4	0.39	0.42	0.32	0.36	0.4	0.43	0.326	0.36	0.4	0.43	0.27	0.3	0.34	0.4
Reject?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Anderson-Darling																
Statistic	0.4439				0.4703				0.5227				0.5401			
P-Value	-				-				-				-			
A	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01
Critical Value	1.93	2.5	3.29	3.91	1.93	2.5	3.29	3.91	1.93	2.5	3.29	3.91	1.93	2.5	3.29	3.91
Reject?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Chi-Squared																
Statistic	0.1877				2.1326				1.2596				0.4733			
P-Value	0.6648				0.1442				0.2617				0.4915			
A	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01
Critical Value	2.71	3.8	5.41	6.63	2.71	3.84	5.41	6.63	2.71	3.84	5.41	6.63	2.71	3.84	5.41	6.63
Reject?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No

sources are displayed in input section of table 3. In addition, by running the proposed configurations in the developed simulation models, we studied the feasibility of these configurations and listed their results in the output section of table 3 for both workflows. The output section consists of the daily labor cost, the number of hours it takes to produce 100 trays, and the number of grafted trays produced.

Table 3 summarizes the labor cost and time to meet the production of 100 trays for each configuration. This table starts by elaborating on the configurations that were suggested by the simulation-based optimization, the scientist, and the manager of a grafting facility. The simulation-based optimization assigned a total of 17 workers for each workflow. With the in-line workflow, 9 of the workers are highly skilled, 5 are moderately skilled, and 3 are low skilled. For the cellular workflow, these numbers change to 7 highly skilled and 10 moderately skilled workers. The plant scientist assigned a total of 19 workers for both workflows, where with the in-line workflow, 17 of the workers are moderately

skilled, and 2 are highly skilled. For the cellular workflow, all 19 workers are moderately skilled. The scientist estimated the cost performance of workers in all skill levels and found that workers with moderate levels of expertise have the highest cost performance and therefore, she assigned a set of 19 workers of moderate skill level for the cellular workflow. As for the in-line workflow, she kept the number of workers the same and allocated workers with different levels of expertise different tasks. Finally, the manager assigned 22 workers for in-line workflow, where 10 of them are highly skilled, 7 moderately skilled, and 5 of them are low skilled. And for cellular, there are 6 workers from each of the three skill levels.

Following the suggested allocations, multiple simulation experiments were conducted to estimate the production capacity and cost of each configuration. Table 3 displays the daily labor cost, hourly production length, and total grafted trays produced solutions provided from all information sources (i.e., simulation-based optimization, scientist, and

Table 3. Allocation of workers by all three information sources (i.e., simulation based optimization, the scientist, and the manager) and the daily cost, length of shift and number of produced trays for each configuration.

			Input - Number of workers					Output		
			Scion Cutter	Root-stock Cutter	Root-stock Clipper	Joining	Total	Daily Labor Cost (\$)	Length (h)	Trays
In-line	Simulation Based Optimization	High	2	3	3	9	17	1,516	7.96	100
		Middle								
		Low								
	Scientist	High	2	2	4	11	19	1,532	7.83	100
		Middle								
		Low								
Cellular	Simulation Based Optimization	High	2	5	5	10	22	1,904	7.90	100
		Middle								
		Low								
	Scientist	High	2	7	10	17	17	1,540	8	100
		Middle								
		Low								
Cellular	Simulation Based Optimization	High	2	6	6	18	18	1,550	8.2	100
		Middle								
		Low								
	Scientist	High	2	6	6	18	18	1,550	8.2	100
		Middle								
		Low								

manager) for both workflows. All suggested configurations are feasible and produce 100 trays per day as required. The solution provided from the simulation-based optimization has the lowest daily cost (i.e., daily labor cost of \$1,516 for in-line and \$1,540 for cellular) for both workflows. The second lowest daily labor cost is achieved by the scientist and manager in workflows 1 and 2, respectively. While the scientist kept the same number of workers for both workflows and increased the overall skill level of the workers, the hourly production length increased due to the longer rework process of the cellular workflow.

Although for the first workflow (i.e., in-line workflow) all the proposed labor allocations are capable of producing more than 100 trays of grafted plants within 8 hours, the additional trays can be considered as waste. To this end, we calculate the labor cost per plant and time per plant using 100 trays per day. Figure 5 compares the cost and time per plant for all the proposed configurations.

In both workflows, the simulation-based optimization setting has the lowest time (i.e., 0.346 for inline and 0.340 min for cellular workflows, respectively) and cost per plant (i.e., \$0.063 for in-line and \$0.064 cellular workflows, respectively). For the in-line workflow, the scientist's solution comes in second place by 0.038 min/plant and \$0.065 /plant, while for the cellular workflow the manager's solution has second place with 0.036 min/plant and \$0.065 /plant.

Although simulation-based optimization claims the optimal/near optimal solution for both workflows, it is important to show that the solutions are statistically different compared to each other. To do so, we first checked the normality of the results, because it is a strong requirement for running further testing regarding the difference within these configurations. While having more than 30 replications of simulation (i.e.,

100 replications in this article) can be considered as a large enough sample size to justify normality, we step further and use normal probability plots to show this. Figure 6 displays the normal distribution aligned to the histogram alongside of the normal probability plot of different settings in both workflows.

As mentioned, figure 6 studies the normality of the results. All six normal probability plots in figure 6, indicate a strong linear relationship between sample data and sample percentile, which supports normality. After displaying normality, Analysis of Variance (ANOVA) is implemented to compare the mean time per plant of all three information sources for each workflow. Table 4 summarizes the ANOVA results.

The hypothesis for these analysis is the equality of sample means in each workflow. To find the different group/groups, the Tukey's studentized range test is conducted to provide a pairwise comparison within the mentioned groups for each workflow. Table 5 summarizes the results of Tukey's studentized range test.

Given the results shown in table 5, we can claim that at $\alpha=0.05$ statistical significance, these configurations have resulted in three different costs per plant for each workflow.

As mentioned in the introduction section, grafting is labor intensive, however this intensiveness faces a higher labor rate in the United States compared to other countries (Calvin and Martin, 2010). and makes up to 42% of the total cost in grafting. Thus, labor costs directly affect the cost of grafted plants and limit their competitiveness against non-grafted ones in the market. In other words, high labor costs may result in the loss of market share for small nurseries. Lower labor costs may lead to higher marginal profits for growers

Figure 5. Time and cost per plant for the labor mix of all three information sources' suggested solutions in both workflows.

Figure 6. Normal probability plots for the simulated results for all the configurations suggested by information sources.

and increase their bargaining power. In this article we successfully decreased labor costs by 20% and 0.6% for in-line and cellular workflows, respectively, via introducing a framework which can allocate labor to different tasks.

Results reveal that by choosing the in-line workflow, the grafting company considered in this study will be able to produce grafted plants with a lower cost. In cases where the solutions of both workflows are equal, it is better to decide the workflow based on the level of expertise of the workers. Generally, if workers have a high or moderate level of expertise, it is better to follow the cellular workflow, whereas lower skilled workers are easier to implement in the in-line workflow. These suggestions rely on the fact that, the cellular workflow encourages competition between workers for higher efficiency and the in-line workflow helps the workers learn faster.

Two key issues (i.e., globality of optimum solution and scalability of the proposed framework) need further discussions, which are addressed in the following sections.

SOLUTION QUALITY

An important aspect of any optimization framework is the quality of the solution (i.e., how close the solution is to the

Table 5. Tukey's studentized range test to identify the groups with different cost per plant.

Workflow	Tukey Grouping	Mean	N	Information Source
In-line	A	0.063	100	Simulation-based optimization
	B	0.064	100	Scientist
	C	0.079	100	Manager
Cellular	a	0.064	100	Simulation-based optimization
	b	0.068	100	Scientist
	c	0.065	100	Manager

global optimum). As mentioned, the proposed framework utilizes both simulation and optimization models in order to find the optimal resource allocation. The question here is whether the framework is capable of finding the global optimum. Although it cannot be proved that the proposed framework guarantees a global optimal solution, we have conducted some analyses to demonstrate the solution quality of this approach. In particular, we have developed a series of experiments based on the "optimize then simulate" approach where the optimal solution of the optimization model is calculated and then those results are implemented in the simulation model to see if they can supply the pre-defined demand of 100 trays in the more realistic environment of

Table 4. Results of ANOVA in order to test the equality of estimated means for cost per plant in each workflow.

Workflow	Source	DF	Sum of Squares	Mean Square	F-Value	P-value
1	Model	2	0.01678752	0.00839376	65849	<0.0001
	Error	297	0.00000381	0.00000001		
	Corrected Total	299	0.01679133			
2	Model	2	0.00328509	0.00164254	555.84	<0.0001
	Error	297	0.00087766	0.00000296		
	Corrected Total	299	0.00416275			

simulation. If the suggested results cannot meet the requirement of producing 100 trays, the optimization model will be forced to utilize additional worker/s. This enforcement is modelled by increasing the right side of equation 6. Figure 7 demonstrates the results.

Figure 7 displays results of the “optimize then simulate” approach. By only running the optimization model, the solution shows that by utilizing 16 workers, it is expected to achieve the production of 100 trays. But the implemented resource allocation in the simulation model could only produce 75 trays per day. Noting that the suggested resource allocation cannot attain our goal, we push equation 6 to be greater than 16. The optimal solution for the upgraded optimization model suggests that using 17 workers, the production capacity will increase to 102 trays per day, while the simulation displays a daily production of 95 trays per day. Similarly, we push the right side of equation 6 to 17. This time the optimum results promised a production capacity of 105 trays per day and the simulation confirmed that this resource allocation of 18 workers can produce the targeted 100 trays per day. Meanwhile, the optimal solution suggested by the proposed framework is 17 workers which has lower labor costs (i.e., \$17,000 less annual labor cost). Although we have shown the capability of our proposed framework, we still need to address the scalability issue.

Scalability

Scalability is an important aspect of any systematic approach. Previously, it has been shown that the proposed framework is capable of finding an optimal/near optimal solution for a small-scale nursery (i.e., fewer than 20 workers). Here, we demonstrate that the proposed framework is scalable for larger facilities. To this end, we conducted a set of experiments to analyze the impact of scale of the problem on solution time. In each experiment, we increased the targeted demand by 100 trays per day and measured the solution time.

As shown in figure 8, the computation time increases by 7 s (i.e., less than 8%) while the scale of the problem has increased tenfold. The polynomial relation between the solution time and the scale of the problem supports the scalability of our proposed framework.

CONCLUSIONS AND FUTURE WORK

In this article, a simulation-based optimization framework has been proposed to support the resource planning and

Figure 7. The solution of “optimize then simulate” approach.

Figure 8. Impact of problem scale on computation time.

development of grafting facilities. In the proposed framework, discrete event simulation models were developed to mimic the two main labor allocation workflows. Optimization models based on the framework were formulated to find the optimal/ near optimal labor allocation for grafting operations. Finally, the results obtained by the proposed simulation based optimization framework were compared to those by the human experts, where a potential annual saving [i.e., (manager teams daily cost – simulation based optimization team daily cost) × number of working days per year] of \$97,388 (20%) for the in-line workflow and a saving of \$2,510 (0.6%) for the cellular workflow were demonstrated.

A possible major extension to the proposed framework is the development of a learning algorithm, which can mimic the dynamic learning behaviour of the workers. Such an algorithm would help predict the performance of the workers and improve the efficiency of the system.

ACKNOWLEDGEMENT

This work has been supported by U.S. Department of Agriculture (USDA) – National institute of food and agriculture under project number 2016-51181-25404.

REFERENCE

- Allen, S. J., & Schuster, E. W. (2004). Controlling the risk for an agricultural harvest. *Manufacturing Service Operations Manag.*, 6(3), 225-236. <https://doi.org/10.1287/msom.1040.0035>
- Andersson, M., & Olsson, G. (1998). A simulation based decision support approach for operational capacity planning in a customer order driven assembly line. *Proc. 30th Conf. on Winter Simulation* (pp. 935-942). Piscataway, NJ: IEEE Computer Society Press.
- Better, M., Glover, F., & Kochenberge, G. (2015). Simulation optimization: Improving decisions under uncertainty. *Breakthroughs in Decision Sci. Risk Analysis*, 18-59.
- Bohle, C., Maturana, S., & Vera, J. (2010). A robust optimization approach to wine grape harvesting scheduling. *European J. Operational Res.*, 200(1), 245-252. <https://doi.org/10.1016/j.ejor.2008.12.003>
- Calvin, L., & Martin, P. (2010). *The U.S. produce industry and labor: Facing the future in a global economy. Economic Research Report-Economic Research Service, USDA*, (106), Washington, DC
- Chiu, Y.-C., Chang, M.-Y., Wu, G.-J., & Chen, C. C. (2008). Development of an automatic outward-feature properties measurement system for grafted tomato seedlings. *Appl. Eng. Agric.*, 24(1), 101-113. <https://doi.org/10.13031/2013.24151>

- Creighton, D., & Nahavandi, S. (2003). Application of discrete event simulation for robust system design of a melt facility. *Robotics Comput.-Integrated Manufacturing*, 19(6), 469-477. [https://doi.org/10.1016/S0736-5845\(03\)00057-7](https://doi.org/10.1016/S0736-5845(03)00057-7)
- Ferrer, J.-C., Mac Cawley, A., Maturana, S., Toloza, S., & Vera, J. (2008). An optimization approach for scheduling wine grape harvest operations. *Int. J. Prod. Economics*, 112(2), 985-999. <https://doi.org/10.1016/j.ijpe.2007.05.020>
- Fishman, G. (2013). *Discrete-event simulation: modeling, programming, and analysis*. New York, New York: Springer Science & Business Media.
- Gosavi, A. (2003). *Simulation-based optimization. parametric optimization techniques and reinforcement learning*. Boston, MA: Springer. <https://doi.org/10.1007/978-1-4757-3766-0>
- Guan, W., Zhao, X., Hassell, R., & Thies, J. (2012). Defense mechanisms involved in disease resistance of grafted vegetables. *HortSci.*, 47(2), 164-170.
- Kazaz, B. (2004). Production planning under yield and demand uncertainty with yield-dependent cost and price. *Manufacturing Service Operations Manag.*, 6(3), 209-224. <https://doi.org/10.1287/msom.1030.0024>
- Kelton, D. W., Sadowski, R. P., & Sadowski, D. A. (2002). *Simulation with arena*. New York, NY: McGraw-Hill.
- Kleijnen, J. P. (2015). *Design and analysis of simulation experiments* (Vol. 20). Springer Int. Publ. <https://doi.org/10.1007/978-3-319-18087-8>
- Kubota, C., McClure, M. A., Kokalis-Burelle, N., Bausher, M. G., & Rosskopf, E. N. (2008). Vegetable grafting: History, use, and current technology status in North America. *HortSci.*, 43(6), 1664-1669.
- Kubota, C., Meng, C., Son, Y.-J., Lewis, M., Spalholz, H., & Tronstad, R. (2017). Horticultural, systems-engineering and economic evaluations of short-term plant storage techniques as a labor management tool for vegetable grafting nurseries. *PLoS One*, 12(2), 1-17. <https://doi.org/10.1371/journal.pone.0170614>
- Laguna, M. (1997). Optimization of complex systems with OptQuest. *A White Paper from OptTek Systems, Inc.*, 1-13. Boulder, Colorado
- Lee, J.-M., Kubota, C., Tsao, S. J., Bie, Z., Echevarria, P. H., Morra, L., & Oda, M. (2010). Current status of vegetable grafting: Diffusion, grafting techniques, automation. *Scientia Hortic.*, 127(2), 93-105. <https://doi.org/10.1016/j.scienta.2010.08.003>
- Louws, F. J., Rivard, C. L., & Kubota, C. (2010). Grafting fruiting vegetables to manage soil borne pathogens, foliar pathogens, arthropods and weeds. *Scientia Hortic.*, 127(2), 127-146. <https://doi.org/10.1016/j.scienta.2010.09.023>
- McCorkle, D., Dudensing, R., Hanselka, D., & Hellman, E. (2016). Evaluating the potential for precision mechanization in US wine grape production. *Proc. Agricultural and Applied Economics Association Annual Meeting* (pp. 1-22). Milwaukee, WI: Agricultural and Applied Economics Association.
- Meng, C., Xu, D., Son, Y.-J., & Kubota, C. (2012). Simulation-based economic feasibility analysis of grafting technology for propagation operation. *Proc. IIE Annual Conf.* (pp. 1-10). Peachtree Corners, GA: Institute of Industrial and Systems Engineers.
- Meng, C., Xu, D., Son, Y.-J., Kubota, C., Lewis, M., & Tronstad, R. (2014). An integrated simulation and AHP approach to vegetable grafting operation design. *Comput. Electron. Agric.*, 102, 73-84. <https://doi.org/10.1016/j.compag.2014.01.007>
- Parthanadee, P., & Buddhakulsomsiri, J. (2010). Simulation modeling and analysis for production scheduling using real-time dispatching rules: A case study in canned fruit industry. *Comput. Electron. Agric.*, 70(1), 245-255. <https://doi.org/10.1016/j.compag.2009.11.002>
- Rasouli, M., Namdari, M., & Mousavi-Avval, H. (2014). Modeling and analysis of energy efficiency in grape production of Iran. *J. Agric. Technol.*, 10, 517-532.
- Rogers, P. (2002). Optimum-seeking simulation in the design and control of manufacturing systems: Experience with OptQuest for Arena. *Proc. Simulation Conf.* (pp. 1142-1150). Piscataway, NJ: IEEE. <https://doi.org/10.1109/WSC.2002.1166371>
- Schwarz, D., & Colla, G. (2010). Amelioration of heavy metal and nutrient stress in fruit vegetables by grafting. *Scientia Hortic.*, 127(2), 165-161.
- Schwarz, D., Rouphael, Y., Colla, G., & Venema, J. H. (2010). Grafting as a tool to improve tolerance of vegetables to abiotic stresses: Thermal stress, water stress and organic pollutants. *Scientia Hortic.*, 127(2), 162-171. <https://doi.org/10.1016/j.scienta.2010.09.016>
- Soto-Silva, W. E., Nadal-Roig, E., Gonzalez-Araya, M. C., & Pla-Aragones, L. M. (2016). Operational research models applied to the fresh fruit supply chain. *European J. Operational Res.*, 251(2), 345-355. <https://doi.org/10.1016/j.ejor.2015.08.046>
- Stoddard, F. L., & McCully, M. E. (1979). Histology of the development of the graft union in pea roots. *Canadian J. Botany*, 57(14), 1486-1501. <https://doi.org/10.1139/b79-185>
- Wan, J., & Li, L. (2008). Simulation for constrained optimization of inventory system by using arena and OptQuest. *Proc. Int. Conf. on Computer Science and Software Engineering* (pp. 202-205). Piscataway, NJ: IEEE.